Национальная академия наук Беларуси

Объединенный институт проблем информатики
Национальной академии наук Беларуси

ПЕРВАЯ ВЫСТАВКА-ФОРУМ IT-АКАДЕМГРАДА «ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В БЕЛАРУСИ»

13–14 октября 2022 года, Минск

Доклады

Минск
ОИПИ НАН Беларуси
2022

:

УДК 004(470+476)(061.3)

Первая выставка-форум IT-академграда «Искусственный интеллект в Беларуси» : сборник докладов, Минск, 13–14 октября 2022 г. – Минск : ОИПИ НАН Беларуси, 2022. – 122 с.

ISBN 978-985-7198-13-9.

В сборнике представлены доклады представителей государственных органов, учреждений образования, ученых и специалистов научно-исследовательских организаций, занимающихся исследованиями в области применения методов искусственного интеллекта.

Адресуется исследователям, практическим работникам и широкому кругу читателей.

Доклады, вошедшие в настоящий сборник, представлены в авторской редакции.

Печатается по решению программного комитета Первой выставки-форума IT-академграда «Искусственный интеллект в Беларуси».

Научное издание

ПЕРВАЯ ВЫСТАВКА-ФОРУМ
IT-АКАДЕМГРАДА
«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ
В БЕЛАРУСИ»

Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/274 от 04.04.2014.
Ул. Сурганова, 6, 220012, Минск.

ISBN 978-985-7198-13-9

УДК 004.912.161.3

КОМПЛЕКС СРОДКАЎ РЭАЛІЗАЦЫІ ЗАДАЧ ШТУЧНАГА ІНТЭЛЕКТУ ДЛЯ БЕЛАРУСКАЙ МОВЫ

Ю. С. Гецэвіч, Я. С. Зяноўка, А. С. Трафімаў, А. А. Бакуновіч, Д. І. Латышэвіч, А. Я. Драгун, М. М. Слесарава, М. С. Тукай
Аб’яднаны інстытут праблем інфарматыкі НАН Беларусі, Мінск

Уводзіны. Хуткае развіцце ІТ-сферы спрыяе новым дасягненням і імкненню ўдасканальваць, мадэрнізаваць і аўтаматызаваць сучаснае грамадства. Праблема штучнага інтэлекту (ШІ), якая закранае такія сферы дзейнасці, як робататэхніка, машыннае навучанне, віртуальная рэальнасць, апрацоўка вялікіх аб’ёмаў даных і інш., пашырае навуковыя даследванні і садзейнічае развіццю перспектыўных практычных рэалізацый. Асноўныя тэндэнцыі развіцця ШІ, якія склаліся ў цяперашні час у сусветным навуковым асяроддзі, перспектывы далейшага развіцця і важнасць удзелу ў навуковай і даследчай працы, дыктуюць галоўныя прыярытэты і напрамкі дзейнасці па павышэнню эфектыўнасці распрацовак, якія ажыццяўляюцца ў названых сферах дзейнасці на тэарытычным і практычным узроўнях. Праграма сацыяльна-эканамічнага развіцця РБ накіравана на інтэлектуалізацыю вытворчасцяў, распрацоўку, паляпшэнне і распашсюджванне беларускай прадукцыі. Адным з важных пунктаў з’яўляецца заахвочванне распрацовак у галіне перспектыўных метадаў штучнага інтэлекту, накіраваных на стварэнне прынцыпова новай навукова-тэхнічнай прадукцыі, у тым ліку на распрацоўку ўніверсальнага (моцнага) штучнага інтэлекту, здольнага думаць, ўзаемадзейнічаць і адаптавацца да зменлівых умоў. Дадзеная мэта рэалізуецца з дапамогай Міжведамаснага даследчага цэнтру штучнага інтэлекту, які быў створаны ў адпаведнасці з пастановай Бюро Прэзідыума НАН Беларусі № 363 ад 31 жніўня 2015 г. на базе Аб'яднанага інстытута праблем інфарматыкі НАН Беларусі і Інстытута фізіялогіі НАН Беларусі [1]. Цэнтр аб'ядноўвае намаганні спецыялістаў у галіне разнастайных навук для стварэння перадавых і канкурэнтаздольных тэхналогій штучнага інтэлекту і стварае ўмовы для выканання навукова-даследчых праектаў у галіне штучнага інтэлекту, што рэалізуюцца як у рамках дзяржаўных праграм навуковых даследаванняў, так і з прыцягненнем недзяржаўных інвестыцый.

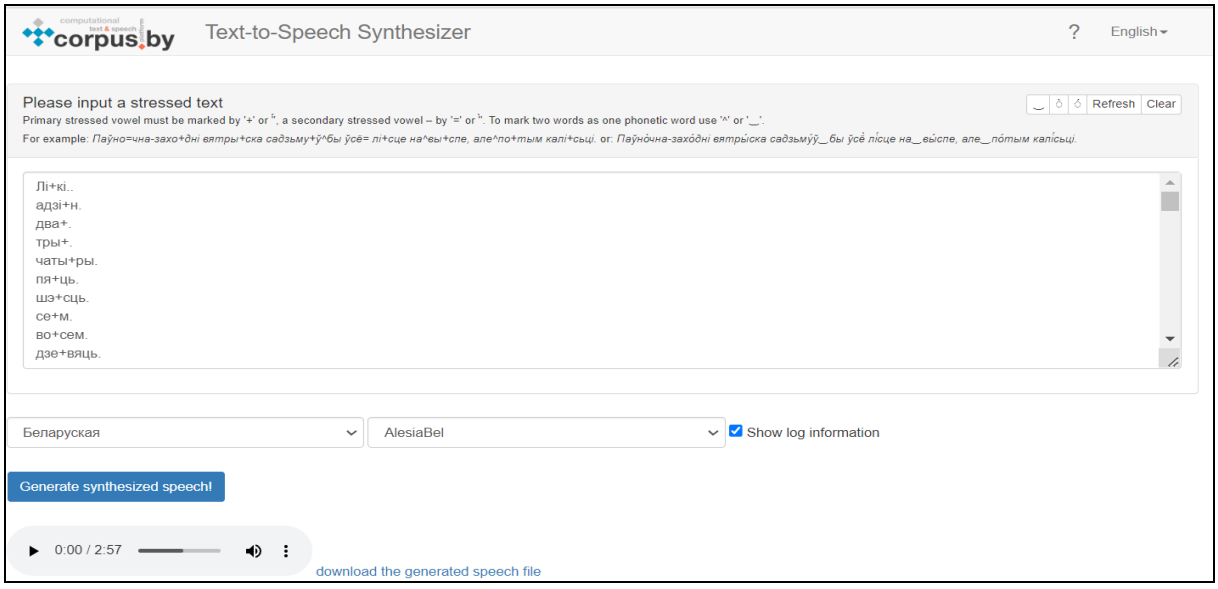
У дадзеным накірунку працуе лабараторыя распазнавання і сінтэзу маўлення АІПІ НАН Беларусі [2], галоўным навуковым прыярытэтам якой з’яўляецца распрацоўка тэорыі і алгарытмаў маўленчага ўзаемадзеяння “чалавек-машына” на аснове глыбокага лінгвістычнага аналізу маўлення і тэксту. Стварэнне прынцыпова новай навукова-тэхнічнай прадукцыі для беларускай мовы, у тым ліку распрацоўка ўніверсальных метадаў, алгарытмаў і прататыпаў, здольных апрацоўваць натуральную мову і маўленне, садзейнічае іх прымяненню ў самых розных галінах навукі, прамысловасці і іншых відаў дзейнасці [3].

Інтэрнэт-платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў. Адной з падзадач штучнага інтэлекту з’яўляецца аўтаматызаваная апрацоўка тэкстаў вялікіх аб’ёмаў, якая атрымала асаблівую актуальнасць. Высокая хуткасць росту колькасці даступнай інфармацыі змушае ўдасканальваць спосабы яе апрацоўкі, рэалізоўваць частковую ці поўную аўтаматызацыю гэтых працэсаў. Прымяненне камп'ютарных тэхналогій для распрацоўкі сістэм, здольных аўтаматычна апрацоўваць электронныя тэксты, дазволіла супрацоўнікам лабараторыі стварыць інтэрнэт-платформу для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў Corpus.by [4], якая дапамагае рашаць мноства задач, звязаных з апрацоўкай электронных тэкстаў і маўленчых

сігналаў. Платформа ўяўляе сабой набор розных сэрвісаў (дакладная колькасць – 74), якія накіраваны на мэтавую аўдыторыю (праграмісты, лінгвісты, філолагі, студэнты, выкладчыкі і г.д.), забяспечваючы просты і ўстойлівы доступ да сродкаў і інструментаў апрацоўкі электроннага тэксту для аналізу, даследавання або аб'яднання такіх набораў даных на беларускай, рускай і англійскай мовах. Прынцып функцыянавання corpus.by заключаецца ў адпаведнасці "уваходныя даныя-выніковыя даныя", дзе карыстальнік уводзіць тэкставую інфармацыю і на выхадзе мае апрацаваныя вынікі. Падыход да распрацоўкі сэрвісаў палягае ў тым, каб карыстальнік мог па ўведзеных тэставых даных запусціць сэрвіс адной кнопкай і азнаёміцца з вынікамі яго працы. Далей карыстальніку прапаноўваецца самастойна выкарыстаць сэрвіс з уведзенымі ўласнымі данымі і выстаўленымі ўласнымі настройкамі.

У аснову кожнага сэрвіса пакладзены асобныя прыёмы і метады фармалізацыі і матэматычнага мадэлявання феноменаў як натуральных, так і штучных (фармальных) моў. Алгарытмізацыя і аўтаматызацыя апрацоўкі фармальна-матэматычных мадэлей, увасабленых у інфармацыйна-тэхналагічных комплексах, рэалізуюцца на аснове фармальных правілаў і граматык, якія адказваюць за лінгвістычную напаўняльнасць сэрвісаў.

Сінтэз маўлення па тэксце на беларускай мове. Акрамя апісанага сэрвіса платформа прадастаўляе інструменты для такенізацыі, марфалагічнага аналізу, агучвання электроннага граматычнага слоўніка, пошуку амонімаў, падліка частотнасці сімвалаў і слоў, праверкі арфаграфіі, сэрвісы па распазнаванні маўлення і эмоцый і многіх іншых. Варта адзначыць і беларускамоўную анлайн-версію сінтэзатара маўлення па тэксце, якая знаходзіцца ў адкрытым доступу [5]. Сістэма рэалізавана на бясплатнай і найбольш распаўсюджанай у Інтэрнэце скрыптавай мове праграмавання PHP. Яна аўтаматычна апрацоўвае ўваходны тэкст на трох мовах (беларускай, рускай і англійскай с беларускім акцэнтам) і фарміруе гукавы файл, які можна праслухаць, спампаваць і захаваць на камп'ютар. Знешні інтэрфейс інтэрнэт-сінтэзатара маўлення паказаны на малюнку 1.

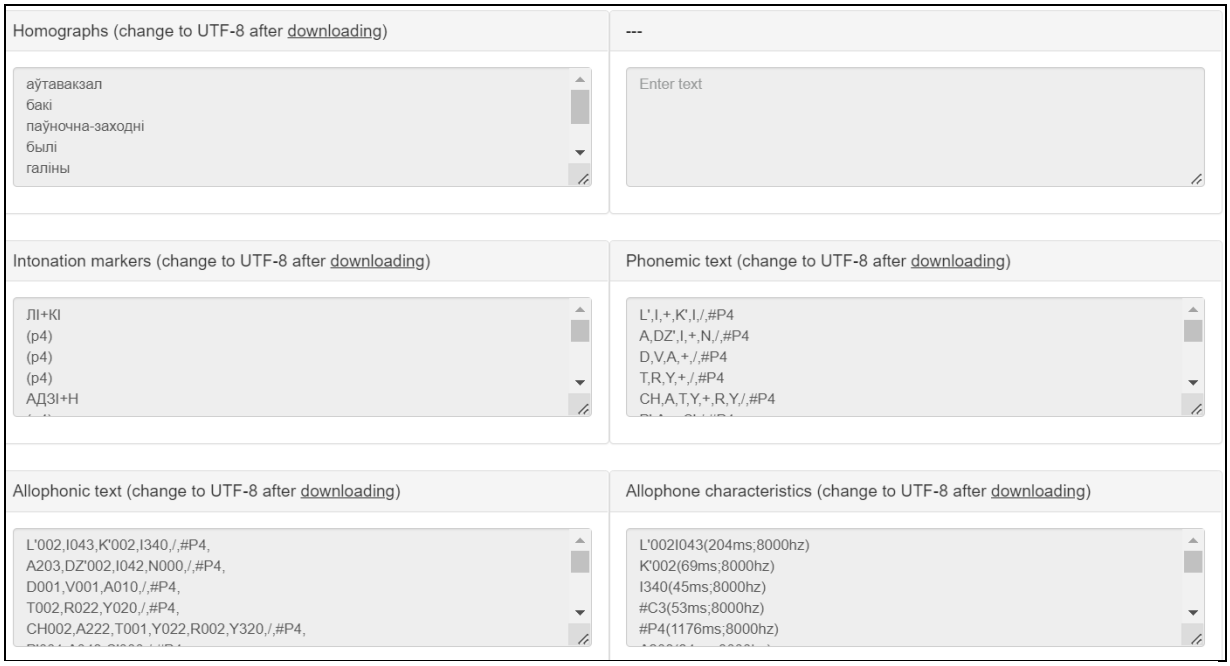


Мал. 1. Знешні інтэрфейс інтэрнэт-версіі сінтэзатара маўлення па тэксце

Лінгвістычная апрацоўка тэксту рэалізавана на высокім узроўні. Невядомыя словы чытаюцца па складах, а словы, напісаныя вялікімі літарамі, вымаўляюцца карэктна. Правільна гучаць складаныя фанетычныя словы: «прыназоўнік+слова», «сло-

ва+часціца». Аднак карэктная апрацоўка спецыфічных токенаў (лікаў, дат, скарачэнняў, адзінак вымярэння, абрэвіятур і інш.) на дадзены момант адсутнічае.

Па ходу пераўтварэння электроннага тэксту ў маўленчы сігнал сінтэзатар генеруе мноства прамежкавых вынікаў. Сярод іх нармалізаваны тэкст, фанемны запіс тэксту, запіс тэксту ў алафонным выглядзеі інш. (мал. 2). Гэтыя вынікі ў сваю чаргу могуць быць выкарыстаны для вырашэння шэрагу іншых камп'ютарна-лінгвістычных задач. Адной з іх з'яўляецца атрыманне транскрыпцыі слоў і цэлых тэкстаў у 4 фарматах: кірылічная, міжнародная (IPA), спрошчаная міжнародная і X-SAMPA. Акно *«Токены»* адлюстроўвае групіраванне сімвалаў па прыкмеце прыналежнасці да аднаго з 5 відаў сімвалаў: *alphabet* (сімвалы мэтавага алфавіту – мовы, абранай для сінтэзавання); *letters* (любыя іншыя літары (не мэтавага алфавіту); *digits* (лічбы); *whitespaces* (прабельныя сімвалы: прабел, перавод радка, табуляцыя і інш.); *character* (іншыя сімвалы).

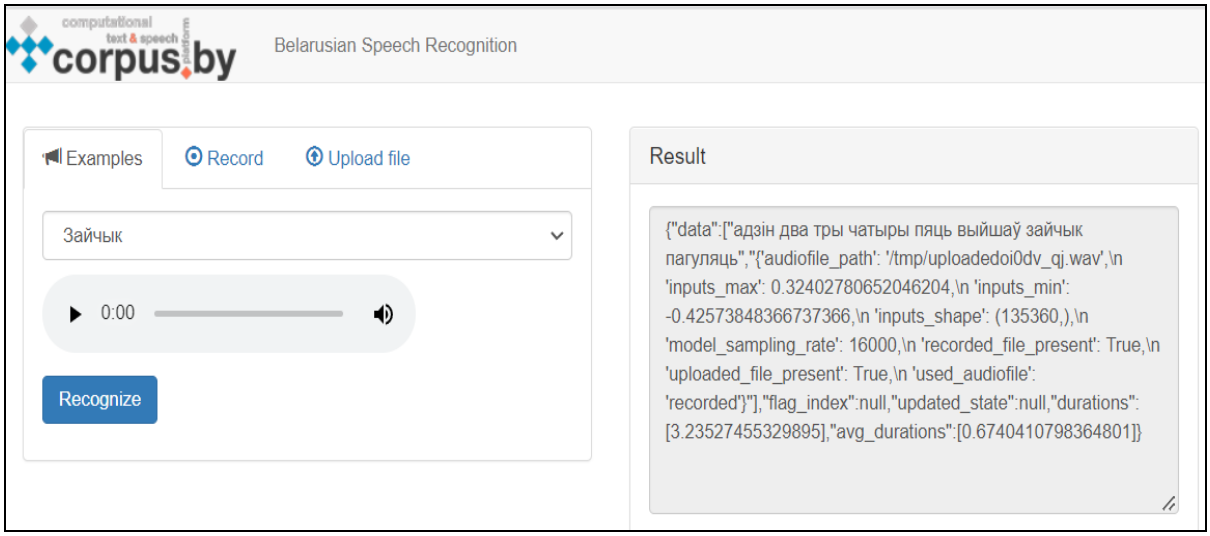


Мал. 2. Прамежкавыя вынікі аналізу ўваходнага тэксту

У *акне «Тэкст»* выводзяцца даныя аб усіх словах, знойдзеных сістэмай у тэксце: вызначаныя часціны мовы і тэгі, якія з'яўляюцца скарочаным абазначэннем марфалагічных прыкмет слова. Акно *«Токены з націскамі»* прадастаўляе спіс слоў з прастаўленымі карыстальнікам націскамі, а *«Невядомыя токены»* – спіс слоў, якія адсутнічаюць у базе даных сістэмы. Інфармацыя аб словах, якія маюць неадназначную пазіцыю націска, прадстаўлена ў палі *«Амографы»* . Спрабуючы вызначыць пазіцыю націска ў слове, сінтэзатар маўлення правярае кожнае слова ўваходнага тэксту на наяўнасць некалькіх спосабаў іх прачытання паводле наяўнай інфармацыі ў слоўніках. Адбываецца пошук слоў з аднолькавым напісаннем і рознымі націскамі. У акне *«Інтанацыйныя пазнакі»* карыстальнік атрымлівае інфармацыю пра інтанацыйную разметку. Вокны *«Фанемны тэкст», «Алафонны тэкст»* прадстаўляюць карыстальніку спіс слоў у фанемным і алафонным выглядах адпаведна. А ў акне *«Характарыстыкі алафонаў»* можна пабачыць працягласці і частоты алафонаў.

Распазнаванне маўлення і гуку. Акрамя сістэм сінтэзу маўлення, лабараторыя таксама актыўна праводзіць даследаванні ў галіне распазнавання беларускага маўлення і галасоў птушак. Распазнаванне маўлення мае вялікія навуковыя перспектывы і шырокія магчымасці прымянення ў шматлікіх сістэмах «чалавек-машына», якія

будуюцца на аснове маўленчых зносін. Распазнаванне беларускага маўлення дасць магчымасць паўнавартаснага развіцця беларускіх тэхнічных навук, у тым ліку робататэхнікі. Асноўныя вынікі прадстаўлены ў двух сэрвісах. Гэта *“Тэматычнае распазнаванне маўлення”* [6] і *“Распазнаванне беларускага маўлення”* [7]. Яны дазваляюць карыстальніку пераўтварыць маўленне ў электронны тэкст анлайн. Інтэрфейс абодвух сэрвісаў падобны (мал.3). Карыстальнік мае магчымасць праслухаць прыклады, якія маюцца ў базе даных, запісаць фанаграмму для распазнавання ці загрузіць гатовы аўдыяфайл з камп’ютара ў фармаце .wav. Пасля запуску апрацоўкі аўдыяфайла сістэма выдае вынік ў выглядзе анлайн-тэкста.

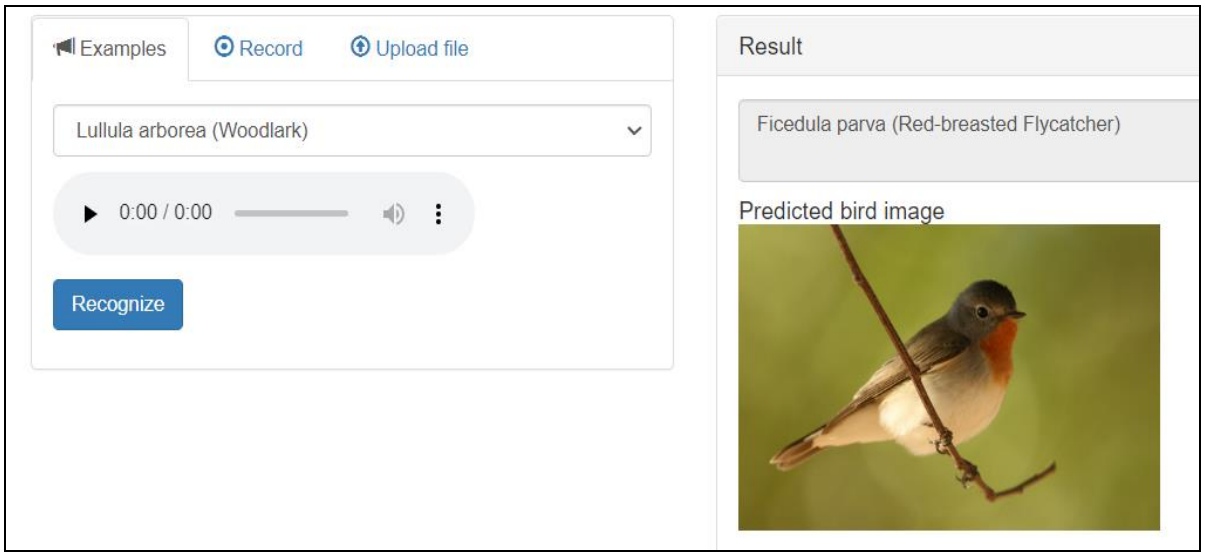


Мал. 3. Інтэрфейс сэрвіса “Распазнаванне беларускага маўлення”

Сэрвіс *“Тэматычнае распазнаванне маўлення”* апрацоўвае фанаграмы маўленчых слоў тэматычных даменаў (вопратка, гарады, лікі, колеры, напрамкі і інш.) памерам ня больш за 20 МВ. На выхадзе сэрвіс выдае распазнаны электронны тэкст. Галоўным недахопам сістэмы з’яўляецца абмежаваная колькасць распазнавання слоў (толькі тыя словы, што ўваходзяць ў базу даных сістэмы). На гэта ўплывае малы памер і варыятыўнасць датасэта, па якім ацэньваецца прататып. Сэрвіс *“Распазнаванне беларускага маўлення”* заснаваны на end-to-end архітэктуры з выкарыстаннем глыбокага навучання. Сістэма распрацавана на базе трэніроўкі і аналізу сабранага датасэту, падчас збора якога ўлічваліся наступныя пункты: высокая варыятыўнасць сабраных аўдыяфайлаў адносна дыктараў (пол, узрост, тэмп маўлення, іншыя асаблівасці вымаўлення), ўмоваў запісаў (розныя мікрафоны, наяўнасць фонавага шуму, інш.). Даденыя асаблівасці дазваляюць навучыць сістэму распазнавання маўлення працаваць ва ўмовах, набліжаных ды тых, з якімі гэтым сістэмам давядзецца працаваць у штодзённым жыцці. Агульная працягласць сабраных аўдыязапісаў на сённяшні дзень складае 987 гадзін (з іх 903 праверана), а ў агучванні тэкстаў прыняло ўдзел 6’160 дыктараў. Гэта першы з падобных датасэтаў настолькі вялікага памеру для беларускай мовы. Сістэма распазнае адвольны тэкст з даволі высокім узроўнем якасці (Test WER = 0.124 або 12.4 %; для параўнання цяперашняе найлепшае значэнне test WER для нямецкага датасэту Common Voice роўнае 5.7).

Ідэнтыфікацыя гукаў жывёл ва ўсёй іх дзікай прыгажосці – гэта вельмі складаная задача. З-за мноства варыяцый гукавых сігналаў розных відаў птушак узнікае праблема іх распазнавання. Зварот да спецыялістаў-арнітолагаў для пошуку і абазначэння той ці іншай пароды патрабуе шмат часу і намаганняў, што прыводзіць да ручной апрацоўкі аўдыяфайлаў. Гэтым даказваецца зацікаўленасць навукоўцаў у стварэнні сістэмы аўта-матызаванага распазнавання галасоў жывёл (на прыкладзе птушак) у цэлым. У рамках

сумеснага праекту з НПЦ НАН Беларусі па біярэсурсах «Распрацоўка тэхналогіі аўтаматызаванага распазнавання галасавых сігналаў жывёл для ажыццяўлення аўтаномнага бесперапыннага маніторынгу рэдкіх відаў з пагрозамі і індыкатарных відаў і стану біяразнастайнасці ў лясных экасістэмах» супрацоўнікі лабараторыі распазнавання і сінтэзу маўлення АІПІ НАН Беларусі стварылі прататып сістэмы аўтаматызаванага распазнавання галасавых сігналаў жывёл для ажыццяўлення аўтаномнага бесперапыннага маніторынгу рэдкіх відаў з пагрозамі, індыкатарных відаў і стану біяразнастайнасці ў лясных экасістэмах [8]. Актуальнасць стварэння такой сістэмы абумоўлена тым, што ўсе існуючыя распрацоўкі па распазнаванню відаў жывёл не падыходзяць для птушак Беларусі, і толькі некаторыя з іх здольныя распазнаваць гукавыя сігналы відаў Еўропы, якія таксама закранаюць дадзены праект.

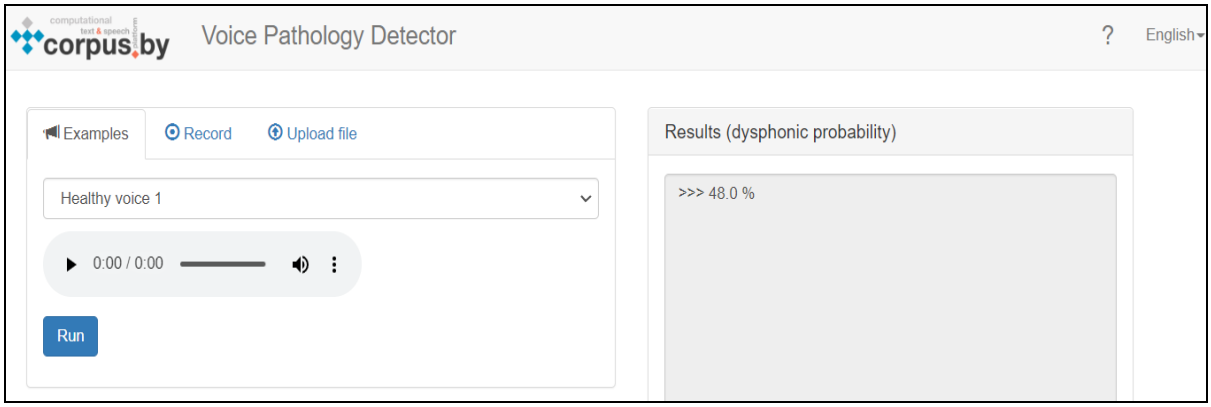


Мал. 4. Інтэрфейс сэрвіса “Распазнаванне галасоў птушак”

Інтэрфейс прыкладання складаецца з двух палей (мал. 4). Першае поле дазваляе выбраць прыклад голаса птушкі з базы даных ва ўкладцы *“Прыклады”*, запісаць свой асобны файл з дапамогай укладкі *“Запісаць”*, ці загрузіць аўдыя з камп’ютара. Акно для праслухоўвання аўдыяфайла прадстаўляе магчымасць спампаваць праслуханы аўдыяфайл. Пасля націскання кнопкі *“Вызначыць”* адбываецца апрацоўка файла і поле *“Вынік”* адлюстроўвае назву прадказанай птушкі з яе выявай. Нягледзячы на тое, што распрацоўка і выкарыстанне аўтаматызаванай сістэмы палягчае працэс распазнавання, сама сістэма мае праблемы дакладнасці распазнавання, якія плануецца выправіць у бліжэйшы час. Праграмнае забяспечанне для вызначэння віду жывёл (на прыкладзе птушак) па галасавых сігналах грунтуецца на матэматычных мадэлях, якія пры недастатковай навучанасці могуць зрабіць вылічальную памылку па вызначэнні неабходнага біялагічнага віду. Таму наступным этапам даследавання з’яўляецца дапрацоўка доследнага ўзору праграмнага забеспячэння баз даных для захавання, выдачы і апрацоўкі галасавых сігналаў жывёл з даступных і ўласных крыніц пасля яго тэсціравання. Прадстаўлены рэсурс прыдатны для шырокага выкарыстання ў навуковых і практыка-арыентаваных даследаваннях, якія патрабуюць апрацоўкі вялікіх аб’ёмаў даных на розных узроўнях.

Адзін з інструментаў па апрацоўцы маўлення, прадстаўленых на платформе, з’яўляецца сэрвіс *“Вызначэнне паталогій голасу”* (мал. 5) [9]. Ён накіраваны на праверку голасу чалавека на наяўнасць паталогіі (захворванняў) для палягчэння дыягнаставання, лячэння і правядзення назіранняў з мэтай аптымізацыі дадзенага працэсу. На сённяшні дзень у гэтай галіне выкарыстоўваюцца праграмна-апаратныя

сродкі. Аднак, невялікі шэраг распрацовак ўяўляе сабой клінічную сістэму аналізу галасоў, здольных не толькі весці вылічэнне характарыстык галасу, але і праводзіць аналіз. Платформа Corpus.by прапаноўвае адзін з падобных сродкаў выяўлення захворванняў маўленчага апарата, якія з'яўляюцца прычынай (або адной з прычын) вакальных характарыстык і параметраў маўлення.



Мал. 5. Інтэрфейс сэрвіса “Вызначэнне паталогій голасу”

На ўваход сэрвісу падаецца аўдыязапіс, які патрабуе праверкі. Па націсканні кнопкі *“Запусціць”* сэрвіс загружае аўдыя на сервер (без захавання), апрацоўвае гэты запіс і вынік адлюстроўваеецца ў выглядзе адзінага параметра – індэкса выяўленнасці дысфаніі (*DSI – Dysphonia Severity Index*). Гэта адсоткак верагоднасці наяўнасці дысфаніі.

Алгарытм разліку DSI рэалізаваны па наступнай формуле:

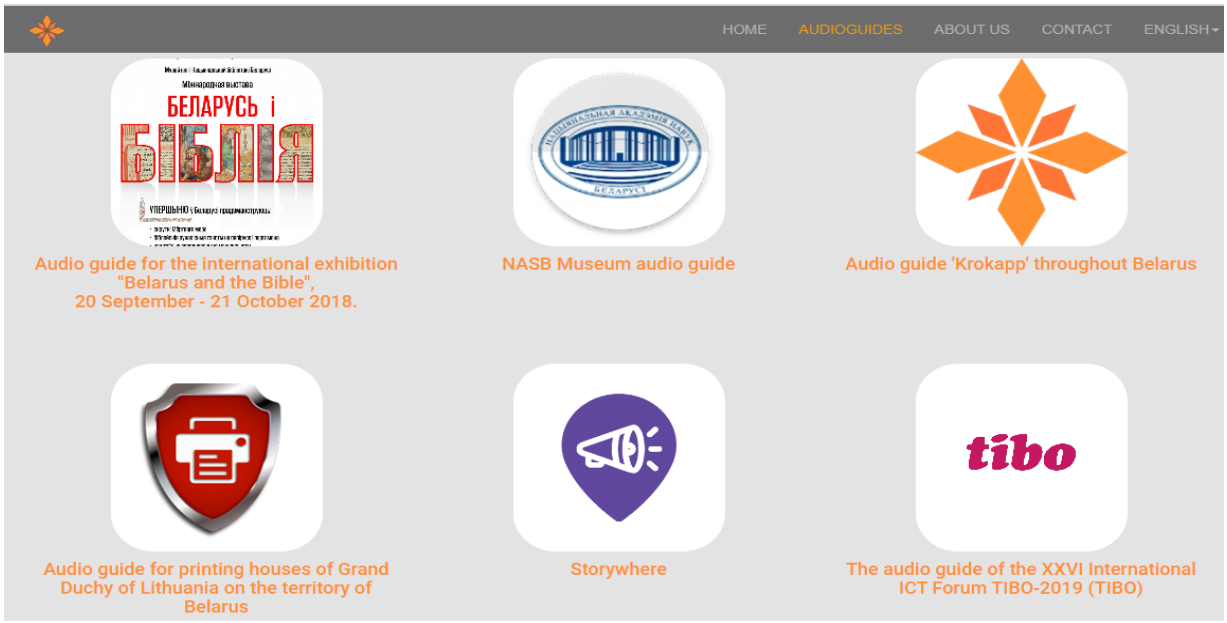
$$DSI = 0.13*ЧМФ+0{,}0053*F0-0.26*I-1.18*Jitter+12.4$$

дзе, I – самая нізкая сіла голасу, F0 – самая высокая частата голасу, ЧМФ – час максімальнай фанацыі, Jitter – нестабільнаць голасу па амплітудзе.

Тое, што на ўваход сэрвіса падаецца файл с запісам гуку, а ня мовы, гаворыць аб тым, што для карыстання сэрвісам ня трэба мець ніякай спецыяльнай падрыхтоўкі. Уваходныя даныя алгарытму: файл з запісам гуку //a// даўжынёй у некалькі секунд. Інтэрфейс складаецца з поля з трымя опцыямі (праслухоўванне прыкладаў, якія маюцца ў сістэме; магчымасць запісу голасу ці загрузкі ўласнага файла з камп’ютара) і тэкставага поля адлюстравання вынікаў. Укладка прыкладаў прадстаўляе набор здаровых голасоў і голасоў з паталогіяй. Адпаведна ў полі вынікаў адлюстроўваецца вылічаны DSI. Укладка загрузкі файлаў дазваляе абраць файл на камп’ютары карыстальніка і загрузіць яго на сервер. Пасля загрузкі файл аўтаматычна апрацоўваецца алгарытмам і ў полі *“Вынікі”* з’яўляецца значэнне DSI.

Платформа аўдыягідаў Krokam. Лічбавыя платформы, заснаваныя на ўзаемадзеянні разнастайных інфармацыйных сістэм, зрабілі больш даступнымі паслугі ў электронным фармаце, у тым ліку і турызме. Распрацаваныя навігацыйныя прыкладанні для смартфонаў, мультымедыя-гіды спрыяюць распаўсюджанню і захаванню гісторыка-культурнай спадчыны нашага народа. На працягу апошніх пяці годоў супрацоўнікі лабараторыі пры падтрымцы спецыялістаў Інстытута гісторыі НАН Беларусі займаюцца распрацоўкамі інтэрактыўных аўдыягідаў па аглядзе культурна-гістарычных славутасцяў Рэспублікі Беларусь. Найбольш значныя распрацоўкі змяшчаюцца на агульнай анлайн-платформе KrokAmm [10], якая прадстаўляе такія плыкладанні, як аўдыягід *“Krokapp”* па ўсёй Беларусі, *аўдыягід па археалагічным музеі "Бярэсце", аўдыягід па міжнароднай выставе "Беларусь і Біблія", аўдыягід па музею НАН Беларусі,аўдыягід 'DRUK VKL' па друкарнях ВКЛ на тэрыторыі Беларусі,*

аўдыягід па XXVI Міжнароднаму форуму па інфармацыйна-камунікацыйных тэхналогіях ТІБО-2019 (TIBO) (мал.6). Усе праграмы асвятляюць розныя культурныя мясціны, гістарычныя падзеі, музейныя выставы і экспанаты, што дазваляе карыстальніку абраць прыкладанне згодна сваім прыярытэтам. Кожны з аўдыягідаў адлюстроўве спіс славутасцей/мясцін/экспанатаў, дае іх короткае тэкставае і/ці агучанае апісанне і іх лакацыю на мапе. Яшчэ адной важнай асаблівасцю кожнага асобнага аўдыягіда з'яўляецца наяўнасць трох версій для розных платформ: web-версія, Android, IOS. Для яе выбару неабходна клікнуць на спасылку на афіцыйным сайце і сістэма аўтаматычна пераадрасуе на выбраную версію для спампоўкі.



Мал. 6. Інтэрфейс анлайн-платформы KrokAmm

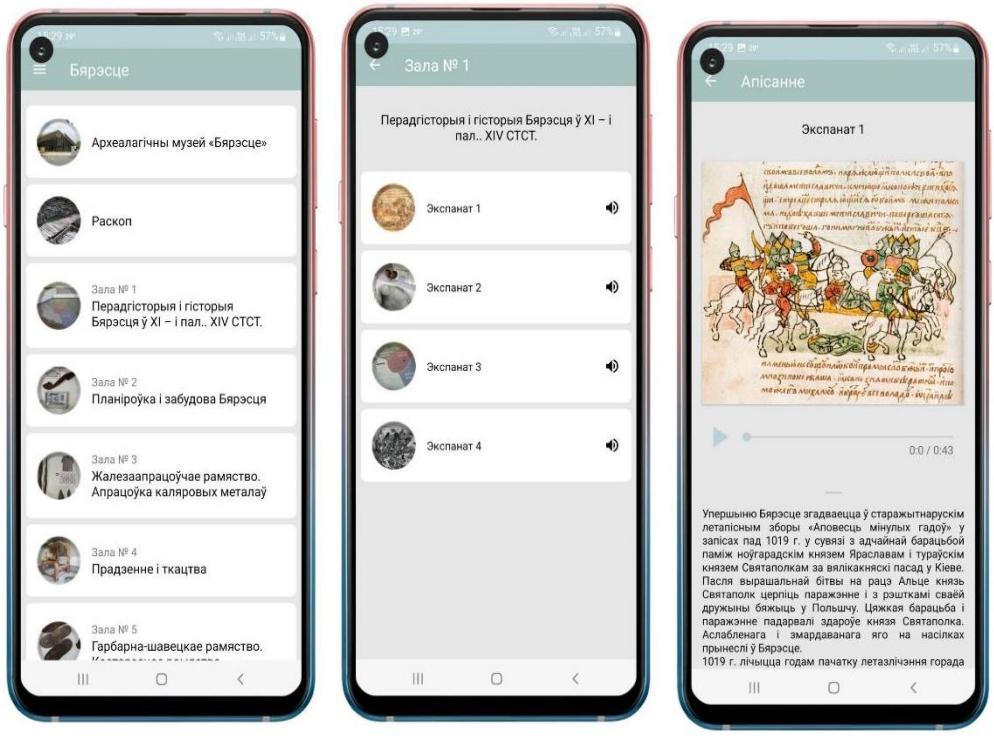
Krokam мае тэхнічныя, эканамічныя і сацыяльныя перавагі. Да тэхнічных адносяцца такія пункты, як зручнасць у карыстанні ўсіх аўдыягідаў; непатрэбнасць у дадатковай тэхнічнай інфраструктуры для сістэмы аўдыягідаў; сістэматызаванае адлюстраванне гарадоў/мясцін/экспанатаў, прадстаўленых у прыкладаннях; хуткае абнаўленне базы даных з дапамогай вэб-дапаможніка ў рэальным часе; лёгкасць рэалізацыі і дабаўлення новых/дадатковых славутасцяў/экспанатаў.

Да эканамічных пераваг адносяцца магчымасць хуткага перабудавання праграмы для выкарыстання ў любым іншым прыкладанні; мінімальны набор рэсурсаў на распрацоўку, як фінансавых, так і чалавечых; якасны сродак дапамогі наяўным экскурсаводам для ахопу вялікай колькасці айчынных і замежных наведвальнікаў праз выкарыстанне аўдыягіда без запрашэння дадатковых экскурсаводаў.

Сацыяльныя перавагі – гэта даступнасць любой версіі (Windows, Android, iOS); падтрымка некалькіх моў (беларуская, англійская, руская, польская, кітайская і інш. у залежнасці ад аўдыягіда), опцыя дабаўлення іншых моў; магчымасць скарыстацца аўдыягідам у зручны час без прывязкі да свабоднага экскурсавода; самастойнае планаванне маршрута асабістай экскурсіі; павышэнне якасці абслугоўвання ў турыстычнай сферы.

Важную ролю ў распрацоўцы аўдыягідаў адыгрывае інфармацыйная частка аўдыягідаў (кантэнт), якая падрыхтавана супольна з супрацоўнікамі музеяў, Інстытута гісторыі НАН Беларусі. Кантэнтная частка праекта ўключае падрыхтоўку электроннага тэксту і выяў аб'екта кантэнт-правайдэрам. Першаснай мовай падрыхтоўкі тэксту з'яўляецца беларуская, далей інфармацыя правяраецца на гістарычную праўдзівасць і

перакладаецца на замежныя мовы з захаваннем асаблівасцей беларускай транслітарацыі. Загрузка даных для мабільных праграм, такіх як выявы і аўдыя- і відэафайлы, а таксама праверка аднаўленняў, адбываюцца не толькі пры дапамозе стандартных бібліятэк Java для Android, але і іншых бібліятэк – Apache, Picasso.



Мал. 7. Android-версія аўдыягіда па археалагічным музеі “Бярэсце”

На платформе можна знайсці аўдыягід для археалагічнага музея ў горадзе Брэсце [11]. Археалагічны музей «Бярэсце» – унікальны археалагічны музей, размешчаны ў горадзе Брэсце, адзіны ў Еўропе музей сярэднявечнага ўсходнеславянскага горада. Аўдыягід для музея распрацаваны з мэтай прадстаўлення ўсіх экспанатаў у даступным фармаце з магчымасцю яго выкарыстання рэальным і віртуальным наведвальнікам падчас знаходжання ў музеі і па-за ім (мал. 7). Прыкладанне прадстаўлена ў выглядзе анлайн-платформы, якая ўключае мабільную праграму (Android, IOS версіі) і вэб-сайт, які рэалізаваны на мове праграмавання Python і базе даных MySQL.

На дадзены момант аўдыягід падтрымлівае 5 моў: беларускую, рускую, англійскую, польскую, кітайскую. У ім прадстаўлены 67 экспанатаў, размешчаныя ў 13 залах, а таксама галоўны экспанат музея – археалагічны раскоп плошчай 1118 квадратных метраў. Кожны экспанат мае невялікае апісанне, фотаздымак і агучанае апісанне на 5 мовах. Для пазіцыянавання па аб’ектах праз аўдыягід у кожнай зале музея пазначаны QR-код, каб карыстальнікі, счытаўшы яго, маглі пазнаёміцца з аб’ектамі візуальным і вербальным спосабамі. Падобным чынам функцыяніруюць усе аўдыягіды, прадстаўленыя на платформе Krokam.

Узаемадзеянне прадстаўленых праграмных прадуктаў дазваляе лабараторыі працаваць у яшчэ адным накрунку: распрацоўка галасавога чат-бота. Саманавучальныя чат-боты, якія змогуць выдаць адказ на любое пытанне карыстальніка/наведвальніка (навігацыя па музеі/залах, гадзіны працы, выбар зручнага маршрута і г. д.) і іх убудаванне ў робатаў-экскурсаводаў з’яўляецца прыярытэтнай задачай на шляху прадстаўлення новых інавацыйных ідэй у культурнай прасторы. Ініцыяцыя камунікацыі чат-бота з чалавекам можа накіравана на прыцягненне ўвагі карыстальніка да

наведвання ці прагляду асобных музеяў/экспанатаў/месцаў, што такім чынам павялічыць зацікаўленасць да культурнай сферы РБ.

Заключэнне. Прыведзеныя праграмныя распрацоўкі лабараторыі распазнавання і сінтэзу маўлення АІПІ НАН Беларусі ўяўляюць сабой зручныя і мнагафункцыянальныя інструменты па апрацоўцы тэкстаў, маўлення і іншых даных на беларускай мове. Яны прадстаўляюць шырокі спект магчымасцей для карыстальнікаў. Устойлівыя Інтэрнэт-распрацоўкі атрымліваюць прымяненне ў разнастаўных Інтэрнэт-праектах, а таксама рэалізуюцца пад мабільныя платформы (рэалізавана для Android, IOS).

Высакаякасны шматмоўны і шматгалосы сінтэз маўлення па тэксце грунтуецца на выкарыстанні алафонных элементаў (усяго парадку 1000 шт.) натуральнага маўлення з максімальна магчымай імітацыяй зададзеных мужчынскіх і жаночых галасоў. Для персаналізацыі камп'ютарнага кланавання натуральнага маўлення было выканана максімальна дакладнае мадэляванне акустычных, фанетычных і прасадычных індывідуальных асаблівасцей голасу і маўлення дыктара з мінімальна магчымымі скажэннямі элементаў кампіляцыі ў працэсе іх запісу, прайгравання і прасадычнай мадыфікацыі. Пабудаваная сістэма распазнавання маўлення для беларускай мовы заснавана на end-to-end архітэктуры з выкарыстаннем глыбокага навучання. Базавыя алгарытмы распазнавання і прыняцця слоўных рашэнняў рэалізуюцца на аснове новага метаду дынамічнага супастаўлення сігналаў, мадыфікаванага для распазнавання слоў злітнага маўлення. Дадзены метад дае магчымасць ажыццявіць дынамічнае выраўноўванне часавых шкал эталоннага апісання слова і яго рэалізацыі ў бягучым маўленні пры невядомых пачатку і канцы слова, якое распазнаецца. Галоўнай вартасцю метаду з'яўляецца вызначэнне верагоднасці прысутнасці слова ў бягучым маўленчым патоку і ацэнкі яго часовага месцазнаходжання ў рэальных умовах рознага роду акустычных памех.

Наяўнасць якасных мадэлей сінтэзу і распазнавання адвольнага маўлення адкрывае для беларускай мовы перспектывы далейшага развіцця больш складаных моўных тэхналогій: галасавы ўвод тэкста, галасавыя дапаможнікі, аўтаматызаванае навучанне беларускай мове, галасавыя чат-боты, і інш. А алічбаванне сферы культуры і распрацоўка аўдыягідаў – гэта першы крок па зборы даных для далейшых вырашэнняў задач штучнага інтэлекту для гіторыка-культурнай спадчыны. Гэта магчыма на прыкладзе прагнозу наведвання канкрэтных мясцін у тым выпадку, калі карыстальнікі цікавяцца іх гісторыяй і культурай. Акрамя таго, дадаткова будзе рэалізавана опцыя збору водгукаў карыстальнікаў пра агляд мясцін гарадоў ці экспанатаў музеяў ў аўтаматычным рэжыме. Забеспячэнне сучаснымі інфармацыйна-камп'ютарнымі сродкамі прыводзіць да фарміравання інавацыйных ідэй, што ўплывае на зацікаўленасць у распрацоўцы новых айчынных прадуктаў у сферы штучнага інтэлекту.

Спіс выкарыстаных крыніц

1. Межведомственный исследовательский центр искусственного интеллекта // ОИПИ НАН Беларуси [Электронный ресурс]. – 2022. Режим доступа : http://uiip.bas-net.by/intellekt/. – Дата доступа : 23.07.2022.

2. Лабараторыя распазнавання і сінтэзу маўлення // [Электронны рэсурс]. – 2022. Рэжым доступу: https://ssrlab.by/. – Дата доступу : 18.02.2020.

3. Мэты дзейнасці // Лабараторыя распазнавання і сінтэзу маўлення АІПІ НАН Беларусі [Электронны рэсурс]. – 2022. Рэжым доступу : https://ssrlab.by/mety-dziejnasci. – Дата доступу : 02.03.2021.

4. Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў corpus.by // [Электронны рэсурс]. – 2019. Рэжым доступу : http://corpus.by. – Дата доступу : 12.07.2021.

5. Сінтэзатар маўлення па тэксце [Электронны рэсурс]. – 2022. Рэжым доступу : http://corpus.by/TextToSpeechSynthesizer/?lang=be. – Дата доступу : 18.03.2022.

6. Тэматычнае распазнаванне маўлення // Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў corpus.by [Электронны рэсурс]. – 2022. Рэжым доступу : https://corpus.by/ThematicSpeechRecognizer/?lang=be. – Дата доступу : 05.06.2022.

7. Распазнаванне беларускага маўлення // Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў corpus.by [Электронны рэсурс]. – 2022. Рэжым доступу : https://corpus.by/BelarusianSpeechRecognition/?lang=be. – Дата доступу : 11.04.2022.

8. Распазнаванне галасоў птушак // Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў corpus.by [Электронны рэсурс]. – 2022. Рэжым доступу : https://corpus.by/BirdSoundsRecognizer/?lang=be. – Дата доступу : 24.06.2022.

9. Вызначэнне паталогій голаса // Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў corpus.by [Электронны рэсурс]. – 2022. Рэжым доступу : https://corpus.by/VoicePathologyDetector/?lang=be. – Дата доступу : 29.05.2022.

10. KROKAM па тутэйшым краі і ваколіцах з вашым асабістым аўдыягідам // [Электронны рэсурс]. – 2021. Рэжым доступу : https://krokam.com/#. – Дата доступу : 01.03.2020.

11. Аўдыягід па археалагічным музеі "Бярэсце" // [Электронны рэсурс]. – 2020. Рэжым доступу : https://krokam.com/biarescie. – Дата доступу : 10.03.2022.