

УДК 004.934

Апрацоўка натуральнай мовы як аснова метадаў і алгарытмаў маўленчага ўзаемадзеяння камп'ютарных сістэм і чалавека

Гецэвіч Юрась Станіслававіч

*загадчык лабараторыі распазнавання і сінтэзу маўлення,
кандыдат тэхнічных навук, дацэнт
E-mail: yuras.hetsevich@newman.bas-net.by*

Лабанаў Барыс Мяфодзьевіч

*галоўны навуковы супрацоўнік лабараторыі распазнавання і сінтэзу маўлення,
доктар тэхнічных навук
E-mail: lobanov@newman.bas-net.by*

Зяноўка Яўгенія Сяргееўна

*малодшы навуковы супрацоўнік лабараторыі распазнавання і сінтэзу маўлення
E-mail: evgeniakacan@gmail.com*

Латышэвіч Давід Іосіфавіч

*малодшы навуковы супрацоўнік лабараторыі распазнавання і сінтэзу маўлення
E-mail: david.latyshевич@gmail.com*

Уводзіны. Лабараторыя распазнавання і сінтэзу маўлення (<https://ssrlab.by/>) была заснавана ў 1974 годзе пад кіраўніцтвам Б. М. Лабанава як аддзел Цэнтральнага навукова-даследчага інстытута сувязі (ЦНДІС), а з 1986 года як лабараторыя Інстытута тэхнічнай кібернетыкі НАН Беларусі. Даследаванні лабараторыі 60-х гадоў былі накіраваны на распрацоўку агульных прынцыпаў аналізу маўленчых сігналаў і вылучэння інфарматыўных прыкмет, якія дазволілі б прадставіць бесперапынны маўленчы сігнал паслядоўнасцю фанетычных прыкмет. У 2000–2010 гадах былі прадстаўлены новая мультыхвалёвая мадэль сінтэзу маўлення па тэксце – «Мультифон» і яе праграмная рэалізацыя, а таксама праграмныя сродкі кланіравання фанетычных і інтанацыйных характарыстык маўлення чалавека – «ФонаКланатар» і «ІнтаКланатар». З пачатку 2010-х гадоў значнай часткай сталі ініцыяваныя Ю. С. Гецэвічам даследаванні, звязаныя з развіццём камп'ютарных метадаў па апрацоўцы тэкставай інфармацыі для сінтэзу і распазнавання маўлення. У 2013 годзе паводле прапановы Б. М. Лабанава Ю. С. Гецэвіч быў прызначаны загадчыкам лабараторыі і з'яўляецца яе кіраўніком у цяперашні час.

На сённяшні дзень галоўны навуковы напрамак дзейнасці лабараторыі – тэорыя распазнавання і сінтэзу маўлення, прымяненне маўленчай сувязі ў чалавека-машынных сістэмах. Акрамя таго, вядуцца праекты ў такіх накірунках, як шматмоўны сінтэз маўлення; распазнаванне дыскрэтнай і злітнай паслядоўнасці слоў; камп'ютарная лінгвістыка; апрацоўка натуральнай мовы; сістэмы з прымяненнем штучнага інтэлекту; распрацоўка вялікіх спецыялізаваных тэкставых і маўленчых баз даных і паралельных карпусоў; распрацоўка пытальна-адказных сістэм, мабільных аўдыягідаў і навігатораў. У лабараторыі працуе 21 чалавек, сярод іх галоўны навуковы супрацоўнік, доктар тэхнічных навук Б. М. Лабанаў; старэйшыя навуковыя супрацоўнікі – кандыдат тэхнічных навук В. А. Захар'еў, кандыдат філалагічных навук М. В. Супрунчук і кандыдат гістарычных навук Н. В. Анофранка, інжынеры-праграмісты, аператары ПЭВМ, малодшыя навуковыя супрацоўнікі і іх стажоры.

Камп'ютарныя тэхналогіі апрацоўкі натуральнай мовы. NLP (Natural Language Processing) – гэта адзін з важных напрамкаў у галіне штучнага інтэлекту, які дазваляе

камп'ютарам разумець і апрацоўваць чалавечую мову. Немалаважным аспектам з'яўляецца яго прымяненне і для беларускай мовы. З мэтай палегчэння працы з тэкстам і аўдыя на дадзенай мове быў створаны анлайн-рэсурс *corpus.by* – платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў. Платформа прапаноўвае карыстальніку набор інструментальных сродкаў (больш за 70 сэрвісаў) па апрацоўцы тэксту, маўлення і іншых даных, якія групуюцца ў тэматычныя дамены для больш зручнага выкарыстання ў асобных сферах дзейнасці. Сэрвісы забяспечваюць просты і ўстойлівы доступ да электроннай апрацоўкі тэксту і маўлення, а таксама да інструментаў для аналізу, выяўлення, даследавання або аб'яднання набораў даных на беларускай, рускай і англійскай мовах. Праца *corpus.by* заснавана на прынцыпе «увядзі свае даныя – атрымай адпаведныя выходныя даныя»: карыстальнік уводзіць тэкставую інфармацыю і атрымлівае апрацаваныя даныя на выхадзе. Платформа прадстаўляе інструменты для такенізацыі, марфалагічнага аналізу, агучвання электроннага граматычнага слоўніка, пошуку амонімаў, падліку частотнасці знакаў і слоў, праверкі арфаграфіі, сінтэзатара маўлення, сэрвісаў распазнавання маўлення, эмоцый і інш. Усе прадукты створаны для вырашэння задач распацоўкі алгарытмаў, рэсурсаў і метадаў інтэрнэт-уводу і інтэрнэт-вываду даных, захавання і сістэматызацыі вялікіх аб'ёмаў інфармацыі. Атрыманыя вынікі апрацоўкі даных могуць быць адаптаваны для шырокага выкарыстання ў прыкладных і практыка-арыентаваных даследаваннях, якія патрабуюць апрацоўкі вялікіх аб'ёмаў даных на розных узроўнях. Тэхналагічны стэк, які выкарыстоўваецца для платформы – PHP (60 %), Python (40 %), MySQL, JavaScript, Html, Css.

На платформе таксама змешчаны такія сэрвісы, як «*Генератар алфавітна-прадметнага паказальніка*», «*Прадказальнік кода УДК*», «*Расшыфроўка УДК*». З іх дапамогай у 2015 годзе былі зроблены значныя крокі ў напрамку развіцця лінгвістычных рэсурсаў і тэхналогій – першае выданне УДК на беларускай мове, якое змяшчае звыш 10 000 класаў асноўных табліц, дапаможныя табліцы і алфавітна-прадметны паказальнік. Праграмны прататып канверціравання арфаграфічнага запісу слова ў фанетычную транскрыпцыю рэалізаваны ў сэрвісе «*Генератар транскрыпцый*». Таксама скампіляваны электронны арфаэпічны слоўнік з дапамогай сэрвіса «*Генератар арфаэпічнага слоўніка*», які з'яўляецца алічбаванай крыніцай «*Арфаэпічнага слоўніка беларускай мовы*» – першага поўнага акадэмічнага слоўніка беларускай мовы, які фіксуе правілы літаратурнага вымаўлення. Платформа таксама прапануе шэраг сэрвісаў для вычыткі беларускамоўных тэкстаў вялікага памеру. У 2018 годзе створана *Методыка і эксперыментальны ўзор аўтаматызаванай вычыткі тэкстаў*, якая дазваляе выявіць арфаграфічныя памылкі ў электронных тэкстах розных тэматык з улікам актуальных правіл правапісу.

Задача перадапрацоўкі электронных тэкстаў на натуральнай мове таксама вырашаецца з дапамогай праграмнага комплексу *NooJ* (<http://www.nooj-association.org/>), які дазваляе распацоўваць сінтаксічныя і марфалагічныя граматыкі ў форме канчатковых аўтаматаў і зручна тэсціраваць іх якасць працы на канкрэтным корпусе тэкстаў. Для гэтых мэт быў непасрэдна распрацаваны беларускі модуль *NooJ*, які ўключае набор скампіляваных тэкстаў, некалькі дэманстрацыйных граматык і набор слоўнікаў (<http://www.nooj4nlp.org/resources/be.zip>).

Маўленчыя тэхналогіі з прымяненнем штучнага інтэлекту. У сучасным свеце тэхналогій сістэмы сінтэзу (СМТ) і распазнавання маўлення (СРМ) адыгрываюць усё больш важную ролю, пашыраючы магчымасці ўзаемадзеяння чалавека з камп'ютарам. Яны забяспечваюць новы ўзровень зручнасці і даступнасці для карыстальнікаў з абмежаванымі магчымасцямі, робяць інфармацыю даступнай больш шырокаму колу людзей і адкрываюць новыя перспектывы ў сферы адукацыі, бізнесу і камунікацыі. Лабараторыя распазнавання і сінтэзу маўлення шчыльна працуе над распацоўкай такіх сістэм на беларускай мове, што спрыяе распаўсюджванню высокатэхналагічных прадуктаў і пашырае выкарыстанне беларускай мовы ў паўсядзённым жыцці і розных сферах дзейнасці.

Сярод беларускамоўных сістэм сінтэзу маўлення варта адзначыць шэраг дадаткаў, якія пераўтвараюць тэкставую інфармацыю ў маўленне. Сярод іх *анлайн-версія сінтэзатара маўлення па тэксце*, якая распрацавана на аснове СМТ «Мультифон-4». Сістэма рэалізавана на

бясplatнай і найбольш распаўсюджанай у інтэрнэце скрыптавай мове праграмавання PHP. Яна аўтаматычна апрацоўвае ўваходны тэкст на трох мовах (беларускай, рускай і англійскай з беларускім акцэнтам) і фарміруе гукавы файл, які можна праслухаць, спампаваць і захаваць на камп'ютар. Лінгвістычная апрацоўка тэксту рэалізавана на высокім узроўні. Падчас пераўтварэння электроннага тэксту ў маўленчы сігнал сінтэзатар генерыруе мноства прамежкавых вынікаў. Сярод іх нармалізаваны тэкст, фанемны запіс тэксту, запіс тэксту ў алафонным выглядзе і інш. Сінтэзатар «*Гаворачая галава*» дазваляе атрымаць візуалізацыю ўведзенага карыстальнікам тэксту ў фармаце відэафайла з аніміраванай галавой. «*Гаворачая галава*» перадае міміку чалавечай галавы і сінтэзаваны гук пры вымаўленні ўваходнага тэксту. Дзякуючы відэаадлюстраванню сэрвіс дазваляе яшчэ больш персаналізаваць уваходны тэкст, паколькі стварае «гаворачыя галовы» на аснове фотаздымкаў рэальна існуючых людзей.

СМТ новага пакалення заснавана на глыбокім навучанні нейронных сетак. У якасці мадэлі для сінтэзу маўлення на беларускай мове была абрана end-to-end мадэль VITS (Variational Inference with adversarial learning for Text-to-Speech). Яна ўяўляе сабой аднаступеньчатую неаўтарэгрэсіўную мадэль, здольную генерыраваць больш натуральны гук у параўнанні з існуючымі двухступеньчатымі мадэлямі, такімі як Tacotron 2, Transformer TTS ці нават Glow-TTS. Выкарыстоўваючы варыятыўную аснову, VITS мадэлюе латэнтную прастору характарыстык маўлення, адлюстроўваючы ўласціваю зменлівасць і нявызначанасць пры генерыраванні маўлення. Для навучання мадэлі абрана база даных, скампіляваная з бібліятэкі аўдыязапісаў на розных мовах, якая размешчана на платформе Mozilla CommonVoice. Для беларускай мовы на сённяшні дзень платформа налічвае 1815 гадзін запісу з 8400 рознымі галасамі ад дыктараў (<https://commonvoice.mozilla.org/be/languages>). Мадэль сінтэзу маўлення навучана з дапамогай Coqui TTS і рэгістратара Weights and Biases на аснове аднаго абранага невядомага дыктара, які начытаў больш за 190 гадзін высокаякаснага маўлення па наяўных тэкставых сказах. Час навучання мадэлі склаў 72 гадзіны. Доступ да мадэлі прадстаўлены ў выглядзе асобнага сэрвіса «*Сінтэзатар беларускага маўлення на тэксце РУ*», які знаходзіцца ў адкрытым доступе на платформе corpus.by (<https://corpus.by/TextToSpeechBelarusianRu/?lang=be>). Прататып характарызуецца простым і зразумелым інтэрфейсам, агучваннем тэкстаў без абмежавання па колькасці слоў і магчымасцю спампавання выніковага аўдыяфайла. Галоўнай перавагай сістэмы з'яўляецца якаснае, зразумелае штучнае маўленне з выразнай інтанацыяй і захаваннем прасадыхных асаблівасцей беларускай літаратурнай мовы.

Акрамя сістэм сінтэзу маўлення лабараторыя актыўна праводзіць даследаванні ў галіне распазнавання беларускага маўлення. Асноўныя вынікі прадстаўлены ў двух сэрвісах. Гэта «*Тэматычнае распазнаванне маўлення*» і «*Распазнаванне беларускага маўлення*». Першы сэрвіс дазваляе карыстальніку пераўтвараць вымаўленыя словы ў электронны тэкст анлайн. Карыстальнік мае магчымасць праслухаць прыклады, якія маюцца ў базе даных, запісаць фанэграму для распазнавання ці загрузіць гатовы аўдыяфайл з камп'ютара ў фармаце .wav. Галоўным недахопам сістэмы з'яўляецца абмежаваная колькасць слоў, што распазнаюцца (толькі тыя словы, што ўваходзяць у базу даных сістэмы). На гэта ўплывае малы памер і варыятыўнасць датасэту, па якім ацэньваецца прататып.

Сэрвіс «*Распазнаванне беларускага маўлення*» заснаваны на end-to-end архітэктурцы wav2vec2 з выкарыстаннем глыбокага машыннага навучання (<https://corpus.by/BelarusianSpeechRecognition/?lang=be>). Яго асаблівасцю з'яўляецца пераднавучанне на корпусе неанатаваных даных для вывучэння спосабаў якаснага вылучэння прыкмет па ўваходным аўдыязапісе. Сістэма распрацавана на базе трэніроўкі і аналізу датасэту, падчас збору якога ўлічваліся такія пункты, як высокая варыятыўнасць сабраных аўдыяфайлаў адносна дыктараў (пол, узрост, тэмп маўлення, іншыя асаблівасці вымаўлення), умовы запісаў (розныя мікрафоны, наяўнасць фонавага шуму, інш.). Дадзеныя асаблівасці дазваляюць навучыць сістэму распазнавання маўлення працаваць ва ўмовах, набліжаных да тых, з якімі гэтым сістэмам давядзецца працаваць у штодзённым жыцці. Агульная працягласць сабраных аўдыязапісаў на сённяшні дзень складае 987 гадзін (з іх 903 правераныя). У агучванні тэкстаў прынялі ўдзел 6160 дыктараў. Гэта першы з падобных датасэтаў такога памеру для беларускай

мовы. Выніковая сістэма распознае адвольны тэкст з даволі высокім узроўнем якасці (Test WER = 0,124 або 12,4 %).

Для беларускамоўных носьбітаў лабараторыя распрацавала інтэрактыўную платформу «Галасавы AI-асістэнт», якая змяшчае шэсць пытална-адказных асістэнтаў жаночага і мужчынскага полу (<https://asistent.io/>). Кожны галасавы асістэнт пабудаваны з выкарыстаннем тэхналогій распазнавання і сінтэзу маўлення, машыннага перакладу і дыялогавых сістэм. Канцэпцыя распрацоўкі грунтуецца на выдачы эфектыўнага і простага ў выкарыстанні механізма прадстаўлення агульнай інфармацыі і вырашэння праблем карыстальнікаў на беларускай мове. AI-асістэнты дазваляюць карыстальнікам вусна ці ў тэкставай форме задаваць пытанні і атрымліваць абгрунтаваны гукавы (надрукаваны) адказ хутка з высокай якасцю і дакладнасцю. Галасавыя памочнікі даступны на афіцыйным сайце платформы, інтэрфейс якой прадстаўлены на беларускай, англійскай, рускай і кітайскай мовах. Там можна абраць асабістага віртуальнага суразмоўца (Vasil, Ales, Alesia, Alena, Boris, Kiryl) і паразмаўляць у інтэрнэце, абраўшы вэб-версію, ці на смартфоне, усталяваўшы дадатак на аперацыйныя сістэмы Android ці iOS. Пытална-адказныя сістэмы таксама даступны ў выглядзе аднайменных чат-ботаў у месэнджары «Тэлеграм». Акрамя таго, можна паразмаўляць з чат-ботаўмі (Vasil, Vasilina), якія вужа накіраваны на розныя сферы дзейнасці (асістэнт агульнага профілю, архітэктар, фінансавы кансультант, гісторык, Францыск Скарына і інш.).

Да тэхналогій аўтаматычнай апрацоўкі і аналізу маўлення таксама адносяцца сістэмы распазнавання галасоў птушак, якія выкарыстоўваюць камп'ютарныя алгарытмы для вызначэння і ідэнтыфікацыі гукаў. Аўтаматычнае распазнаванне галасоў птушак шырока ўжываецца ў навуковых даследаваннях, маніторынгу біяразнастайнасці і ацэнцы экалагічнага стану асяроддзя. Для вырашэння пералічаных задач быў распрацаваны *Інфармацыйна-аналітычны цэнтр бесперапыннага аўтаматызаванага маніторынгу рэдкіх, пагражальных і індыкатарных відаў жывёл (птушак) Рэспублікі Беларусь*. У рамках праекта распрацоўкі інфармацыйна-аналітычнага цэнтра былі выкананы наступныя задачы: стварэнне праграмага асяроддзя для аўтаматызаванага распазнавання галасавых сігналаў жывёл на аснове выкарыстання распрацаванай камп'ютарнай мадэлі навучання; збор гукавых запісаў птушак для кампіляцыі базы даных з даступных і ўласных крыніц; анатаванне галасавых запісаў жывёл у мэтах стварэння і наступнага паляпшэння мадэлі аўтаматызаванага распазнавання; падрыхтоўка мадэлі і алгарытмаў функцыянавання праграмага забеспячэння, заснаваных на глыбокім машынным навучанні скруткавай нейроннай сеткі; распрацоўка і тэсціраванне вэб- і мабільных дадаткаў распазнавання галасавых сігналаў жывёл.

На сённяшні дзень распрацаваны мадэль распазнавання галасоў птушак, алгарытм аўтаматызаванай разметкі галасавых сігналаў і прататып эксперыментальнага праграмага забеспячэння інфармацыйна-аналітычнага цэнтра. Праграмае забеспячэнне інтэгравана ў інфармацыйна-аналітычны цэнтр бесперапыннага маніторынгу і даступна для выкарыстання ў адкрытым доступе ў фармаце вэб-платформы (<https://bird-voice-iac.ssrilab.by/>), Android- і iOS-версій для мабільных тэлефонаў.

Алічбаванне гісторыка-культурнай спадчыны з персанальнымі аўдыягідамі. Лічбавыя платформы, заснаваныя на ўзаемадзеянні разнастайных інфармацыйных сістэм, зрабілі больш даступнымі паслугі ў электронным фармаце, у тым ліку і турызме. На працягу апошніх сямі гадоў супрацоўнікі лабараторыі пры падтрымцы спецыялістаў Інстытута гісторыі НАН Беларусі займаюцца распрацоўкамі інтэрактыўных аўдыягідаў і навігацыйных сістэм па аглядзе культурна-гістарычных славуцасцей і значных мясцін Рэспублікі Беларусь. Распрацоўкі змяшчаюцца на агульнай анлайн-платформе *Krokam* (<https://krokam.com/>), якая прадстаўляе такія дадаткі, як аўдыягід «Крокар» па ўсёй Беларусі, аўдыягід па археалагічным музеі «Бярэсце», аўдыягід па міжнароднай выставе «Беларусь і Біблія», аўдыягід па музеі НАН Беларусі, аўдыягід «DRUK VKL» па друкарнях ВКЛ на тэрыторыі Беларусі, аўдыягід па XXVI Міжнародным форуме па інфармацыйна-камунікацыйных тэхналогіях «ТІВО-2019», аўдыягіды «Лідскі замак» і «Батанічны сад», «Навігатар па рыбных зімавальных ямах». Пералічаныя дадаткі асвятляюць розныя культурныя мясціны, гістарычныя падзеі, музейныя

выставы і экспанаты, што дазваляе карыстальніку абраць найбольш зручны варыянт згодна сваім прыярытэтам. Кожны з аўдыягідаў адлюстроўвае спіс славутасцей/мясцін/экспанатаў, дае іх кароткае тэкставае і (ці) агучанае апісанне, іх лакацыю на карце. Яшчэ адной важнай асаблівасцю кожнага асобнага аўдыягіда з'яўляецца наяўнасць трох версій для розных платформ: вэб-версія, Android, iOS.

Платформа Krokam мае тэхнічныя, эканамічныя і сацыяльныя перавагі. Да тэхнічных адносяцца такія пункты, як зручнасць у карыстанні ўсімі аўдыягідамі; непатрэбнасць у дадатковай тэхнічнай інфраструктуры для сістэмы аўдыягідаў; сістэматызаванае адлюстраванне гарадоў/мясцін/экспанатаў, прадстаўленых у дадатках; хуткае абнаўленне базы даных з дапамогай вэб-дапаможніка ў рэальным часе; лёгкасць рэалізацыі і дабаўлення новых (дадатковых) славутасцей/экспанатаў. Да эканамічных пераваг адносяцца магчымасць хуткага перабудавання аўдыягіда для выкарыстання ў любым іншым дадатку; мінімальны набор рэсурсаў на распрацоўку як фінансавых, так і чалавечых; якасны сродак дапамогі наяўным экскурсаводам для ахопу вялікай колькасці айчынных і замежных наведвальнікаў. Сацыяльныя перавагі – гэта даступнасць любой версіі (Windows, Android, iOS), падтрымка некалькіх моў (беларускай, англійскай, рускай, польскай, кітайскай і іншых у залежнасці ад аўдыягіда), опцыя дабаўлення іншых моў, магчымасць скарыстацца аўдыягідам у зручны час без прывязкі да свабоднага экскурсавода, самастойнае планаванне маршруту асабістай экскурсіі, павышэнне якасці абслугоўвання ў турыстычнай сферы.

Даследаванне прасадыхных характарыстык мовы. Пачынаючы з 2015 года, працы Б. М. Лабанава прысвечаны даследаванню ўласцівасцяў маўленчай інтанацыі, неабходных для ўдасканалення валодання мовай як чалавекам, так і камп'ютарнымі сістэмамі распазнавання і сінтэзу маўлення. Распрацаваны шматмоўныя сістэмы *IntonTrainer*, *Speech Rate Meter*, *Speech Melody Meter* і *Singer VoiceTester*, якія размешчаны на сайце <https://intontrainer.by>. Усе праграмы агульнадаступны для бясплатнай загрузкі і тэсціравання.

Праграмны модуль «IntonTrainer» прызначаны для аналізу, экраннага адлюстравання і супастаўлення інтанацыйных контураў частаты асноўнага тону эталоннай і вымаўленай фраз, а таксама лічбавай ацэнкі іх інтанацыйнага падабенства. Комплекс дазваляе ажыццяўляць першаснае азнаямленне з асноўнымі інтанацыйнымі канструкцыямі маўлення, якія выкарыстоўваюцца ў розных мовах (англійскай, рускай, кітайскай, беларускай), іх супастаўленне, вывучэнне асаблівасцяў ужывання, а таксама іх рэалізацыі ў дыялогу, прозе і вершах. Праграма *Speech Rate Meter* прызначана для ацэнкі комплексу прасадыхных характарыстык тэмпу маўлення, а менавіта: агульнай хуткасці чытання тэксту, хуткасці генерацыі фанематычнага патоку, сярэдняй працягласці міжфразавых паўз, адноснай працягласці «запаўнення» вакальных уставак. *Speech Melody Meter* ацэньвае статыстычныя характарыстыкі мелодыкі маўлення шляхам пабудовы і адлюстравання гістаграм размеркавання частот сустракаемасці дыскрэтных значэнняў частаты асноўнага тону на працяглых інтэрвалах маўлення. Праграма *Singer VoiceTester* выкарыстоўваецца для тэсціравання спеўных здольнасцяў карыстальніка шляхам аналізу статыстычных характарыстык працяглага інтэрвалу (парадку 10 секунд) акустычнага сігналу яго спеваў. Дадатак *Voice EmoMeter* прымяняецца для аналізу і вымярэння ступені эмацыянальнасці голасу чалавека як для запісанага непасрэдна з мікрафона маўленчага фрагмента, так і для запісаных загадзя тэставых аўдыяфайлаў з розных крыніц.

Прасадыхныя характарыстыкі маўлення разам з эмацыянальнасцю маўлення характарызуюць галоўныя ўласцівасці маўленчай дзейнасці асобы, што немалаважна для інтанацыйна-сэнсавага дзялення маўлення, перадачы камунікатыўнага тыпу выказвання, вызначэння асобных эмацыянальных значэнняў. Іх рэалізацыя ў складзе камп'ютарных сістэм з акустычным кампанентам прыбліжае працэс надання натуральнага гучання штучнаму маўленню і абумоўлівае важнасць вывучэння інтанацыйных параметраў на камп'ютарным узроўні.

Акамуляцыя ведаў у галіне штучнага інтэлекту. Хуткае развіццё IT-сферы спрыяе новым дасягненням і імкненню ўдасканальваць, мадэрнізаваць і аўтаматызаваць сучаснае грамадства. Прымяненне штучнага інтэлекту у такіх галінах, як робататэхніка, машыннае навучанне, віртуальная рэальнасць, апрацоўка вялікіх аб'ёмаў даных і інш., пашырае навуковыя даследаванні і садзейнічае развіццю перспектыўнай практычнай рэалізацыі. Укараненне

сістэмных тэхналогій на аснове штучнага інтэлекту патрабуе комплекснага падыходу як міжнароднага досведу і прадуктаў штучнага інтэлекту ад лідараў у адпаведных галінах, так і прадстаўлення айчынных распрацовак для выкарыстання зацікаўленымі на міжнароднай арэне. Для аб'яднання і акумуляцыі перадавых даследаванняў і распрацовак штучнага інтэлекту, створаных у Рэспубліцы Беларусь, а таксама наладжвання камунікацыі і сувязі паміж навуковымі суполкамі лабараторыяй распазнавання і сінтэзу маўлення прадстаўлена «*Платформа штучнага інтэлекту BELAI*». Тэрміновасць, ініцыятыўны, навучальны і інтэграцыйны падыход, удзел усіх зацікаўленых бакоў і правядзенне публічных абмеркаванняў будуць складаць асноўныя накірункі дзейнасці платформы, каб дазволіць Беларусі вырашаць, як яна жадае выкарыстоўваць штучны інтэлект для росквіту і дабрабыту.

Мэта распрацоўкі платформы – стымуляваць дзейнасць у галіне штучнага інтэлекту і выпрацаваць агульныя прынцыпы работы па развіццю штучнага інтэлекту ў Беларусі. Гэта партнёрства, у якім прамысловасць, адукацыйныя і навукова-даследчыя ўстановы, бізнес і зацікаўленая супольнасць імкнуцца паскараць развіццё штучнага інтэлекту і аб'ядноўваць ініцыятывы штучнага інтэлекту ў Беларусі.

З канцэптуальным вэб-прыкладаннем можна азнаёміцца на афіцыйным сайце па спасылцы <https://belai.by/>. Інтэрфейс сайта прадстаўлены на беларускай і рускай мовах. Раздзел *Навіны* інфармуе супольнасць па тэмах, звязаных са штучным інтэлектам, або генерыруе адпаведныя дыскусіі. Раздзел *Даследаванні* ўключае навуковыя публікацыі, артыкулы па пэўных тэмах, можа змяшчаць спасылкі на базы даных са штучным інтэлектам, спасылкі на запісы вэбінараў, інтэрв'ю, вэб-сайты або любы іншы кантэнт, які дапамагае павысіць дасведчанасць у працы і інтарэс удзельнікаў, звязаных са штучным інтэлектам. Раздзел *Канферэнцыі* накіроўвае на галоўныя падзеі ў галіне штучнага інтэлекту Беларусі.

Для пашырэння ведаў беларускага грамадства па сучасных айчынных і сусветных дасягненнях у штучным інтэлекце створаны тэлеграм-канал BelAI як дадатковы інструмент асвятлення галоўных падзей, навін і даследаванняў у сферы штучнага інтэлекту (<https://t.me/belaiplatform>). Распрацоўшчыкі канала выкарыстоўваюць розныя фарматы кантэнту, такія як тэкставыя нататкі, фота, відэа і апытанні, каб падтрымліваць узаемадзеянне з падпісантамі. Тэлеграм-канал публікуе навіны два разы ў дзень і знаёміць карыстальнікаў з наступнымі відамі навін: асвятленне актуальных даследаванняў і распрацовак у сферы штучнага інтэлекту ў свеце і Беларусі; знаёмства з сучаснымі камп'ютарнымі сэрвісамі, праграмамі і прыладамі, якія выкарыстоўваюць тэхналогіі штучнага інтэлекту; адлюстраванне апошніх дасягненняў штучнага інтэлекту ў робататэхніцы, медыцыне, прыборабудаванні, камп'ютарных тэхналогіях; прадстаўленне разнастайных бібліятэк, камп'ютарных інфармацыйных сховішчаў, якія даюць доступ да вялікай колькасці інструментаў з штучным інтэлектам; апісанне статыстычных даных развіцця штучнага інтэлекту ў часовай адлегласці, таксама параўнанне штучнага інтэлекту з чалавечымі магчымасцямі.

Мяркуецца, што платформа штучнага інтэлекту і тэлеграм-канал паспрыяюць знаёмству і аб'яднанню беларускіх кампаній, каманд і індывідуальных распрацоўшчыкаў і практыкаў беларускай супольнасці ў галіне штучнага інтэлекту, а таксама пашыраць веды ў гэтай галіне.

Удасканаленне працы аўтаматызаваных сістэм па тэкстах розных тэматык. У межах задання «Мадэлі, метады, алгарытмы і праграмныя сродкі інтэлектуальнай апрацоўкі, аналізу і распазнавання медыка-біялагічных даных, малюнкаў, маўленчай і тэкставай інфармацыі, распрацоўка на іх аснове інфармацыйных тэхналогій і сістэм медыцынскага і сацыяльнага прызначэння» былі пастаўлены дзве асноўныя падзадачы:

1. Распрацоўка мадэляў паралельных электронных карпусоў маўлення і тэкстаў медыцынскай, сацыяльнай і прававой тэматыкі на рускай, беларускай і англійскай мовах для распрацоўкі і тэсціравання інтэлектуальных сістэм распазнавання і сінтэзу маўлення.

2. Распрацоўка ўніверсальных алгарытмаў апрацоўкі маўлення і тэкстаў па сабраных электронных карпусах маўлення і тэкстаў медыцынскай, сацыяльнай і прававой тэматыкі на рускай, беларускай і англійскай мовах.



Калектыў лабараторыі распазнавання і сінтэзу маўлення



Удзел у II Форуме ІТ-Акадэміі «Штучны інтэлект у Беларусі», 2023 г.



Удзел прадстаўнікоў лабараторыі ў эфіры беларускага радыё «Культура»

Для дасягнення гэтай мэты лабараторыя ажыццявіла апрацоўку тэкстаў прыведзеных тэматык і іх назапашванне для кампіляцыі электронных паралельных карпусоў тэкстаў. Першапачатковым этапам стаў пераклад на беларускую мову кодэксаў законаў, паколькі толькі адзін з іх («Аб культуры») быў прыняты па-беларуску. У чэрвені 2024 года гэта праца была скончана. Яна вялася ў некалькі этапаў. Спачатку кодэкс перакладаліся з дапамогай машыннага перакладчыка на сайце pravo.by, пасля іх вычыталі два рэдактары-лінгвісты і адзін рэдактар-юрыст. У працэсе ручной апрацоўкі вынікаў аўтаматычнага перакладу былі выяўлены разнастайныя недакладнасці, памылкі і хібы, праведзена ўніфікацыя тэрмінаў і ўстойлівых выказаў. Затым тэксты правяраліся з дапамогай карэктарскіх сэрвісаў сайта corpus.by. Далей тэксты з дапамогай сінтэзатара маўлення былі канвертаваны ў аўдыяфайлы кодэксаў, што забяспечыць доступ да гэтай інфармацыі і для людзей са слабым зрокам, а самі сінтэзатары могуць быць пратэставаны на тэкстах юрыдычнай тэматыкі. Акрамя апісанага прыкладнага значэння праект мае таксама адукацыйную і навуковую вартасць. Так, на розных этапах да яго далучаліся студэнты БДУ, якія выконвалі ў яго рамках свае заданні для дыпломных прац ці вытворчай практыкі. Падрабязней аб праекце можна прачытаць тут – <https://ssrllab.by/7804>.

Акрамя таго, былі сфарміраваны карпусы медыцынскай і сацыяльнай тэматык (на аснове паралельных перакладаў матэрыялаў медыцынскіх арганізацый і гісторыка-культурнага сацыяльнага аўдыягіда KrokApp адпаведна). Карпусы прадстаўлены на беларускай, рускай і англійскай мовах у фармаце TXT і NooJ. Назіранні над спалучальнасцю, спецыфікай тэрмінаўтварэння, недакладнасці камп'ютарнага ці ручнога спосабу перакладу асэнсаваны ў шэрагу артыкулаў, падрыхтаваных супрацоўнікамі лабараторыі. Значную карысць у гэтай працы мае лінгвістычны працэсар NooJ. Ён выкарыстоўваўся для складання карпусоў канкрэтных тэматык тэкстаў з граматычнай разметкай і для алгарытмаў (сінтаксічных граматык), якія дапамагалі ўдакладняць значэнні слоў, ліквідаваць іх шматзначнасць у тэкстах.

Заклучэнне. Прыведзеныя праграмныя распрацоўкі лабараторыі распазнавання і сінтэзу маўлення ўяўляюць сабой зручныя і шматфункцыянальныя інструменты па апрацоўцы тэкстаў, маўлення і іншых даных на беларускай мове. Яны прадстаўляюць шырокі спектр магчымасцей для карыстальнікаў, якія дазваляюць выконваць складаныя задачы ў тэарэтычных даследаваннях і прыкладных праграмах. Устойлівыя інтэрнэт-распрацоўкі атрымліваюць прымяненне ў разнастайных інтэрнэт-праектах, а таксама рэалізуюцца пад мабільныя платформы (для Android, iOS).

Распрацоўка платформы corpus.by па апрацоўцы натуральнай мовы спрыяе хуткаму і ўстойліваму доступу да сэрвісаў апрацоўкі тэкстаў і маўлення на розных мовах, разнастайных дадаткаў для аўтаматызацыі працэсаў і алгарытмаў, акумуляванню новых моўных і маўленчых рэсурсаў, што абумоўлена хуткім тэмпам развіцця тэхналогій штучнага інтэлекту.

Наяўнасць якасных мадэлей сінтэзу і распазнавання адвольнага маўлення адкрывае для беларускай мовы перспектывы далейшага развіцця больш складаных моўных тэхналогій: галасавы ўвод тэксту, галасавыя дапаможнікі, аўтаматызаванае навучанне беларускай мове, навігацыйныя сістэмы і інш. Алічбаванне сферы культуры і распрацоўка сучасных аўдыягідаў у першую чаргу накіраваны на прыцягненне ўвагі да гісторыка-культурных каштоўнасцей краіны і павышэнне якасці жыцця насельніцтва. У будучым плануецца інтэграцыя ў аўдыягіды і навігатары галасавых памочнікаў штучнага інтэлекту. Забеспячэнне сучаснымі інфармацыйна-камп'ютарнымі сродкамі прыводзіць да фарміравання інавацыйных ідэй, што ўплывае на зацікаўленасць у распрацоўцы новых айчынных прадуктаў у сферы штучнага інтэлекту. Развіццё штучнага інтэлекту мае вялікі патэнцыял для ўдасканалення навуковых даследаванняў, павышэння эфектыўнасці вытворчасці, стварэння новых бізнес-мадэляў і паляпшэння якасці жыцця ў краіне. Стварэнне агульнай Платформы штучнага інтэлекту Рэспубліцы Беларусь і навінавага тэлеграм-канала BelAI будзе садзейнічаць сумеснаму развіццю айчынных перадавых тэхналогій штучнага інтэлекту. Камбінацыя хуткага доступу да інфармацыі, разнастайнай тэматыкі, інтэрактыўнасці і лёгкасці выкарыстання робіць прадстаўленыя рэсурсы прывабнымі для шырокага кола распрацоўшчыкаў і ўсіх зацікаўленых, дазваляючы супрацоўніцтва, абмен ведамі і вопытам, а таксама стварэнне і ўдасканальванне алгарытмаў, мадэлей і архітэктуры штучнага інтэлекту.