

Объединенный институт проблем информатики
Национальной академии наук Беларуси

XXII Международная
научно-техническая конференция

**РАЗВИТИЕ ИНФОРМАТИЗАЦИИ
И ГОСУДАРСТВЕННОЙ СИСТЕМЫ
НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ**

РИНТИ-2023

16 ноября 2023 г., Минск

Доклады

Минск
ОИПИ НАН Беларуси
2023

Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2023) : доклады XXII Международной научно-технической конференции, Минск, 16 ноября 2023 г. – Минск : ОИПИ НАН Беларуси, 2023. – 400 с. – ISBN 978-985-7198-15-3.

Представлены доклады XXII Международной научно-технической конференции «Развитие информатизации и государственной системы научно-технической информации» (РИНТИ-2023), Минск, 16 ноября 2023 г., в которых рассмотрены концептуальные основы создания Единого республиканского центра организации доступа к мировым электронным информационным ресурсам, результаты научно-методического обеспечения развития информатизации в 2022–2023 гг., роль человеческого капитала в развитии цифровизации и информационного общества в контексте качества образования, состояние и перспективы цифрового развития Республики Беларусь, вопросы непрерывной подготовки кадров в области кибербезопасности для цифровой трансформации и цифрового развития отраслей экономики Беларуси, необходимые условия технологического суверенитета в сфере ИКТ, факты истории создания белорусской вычислительной техники и др.

Рассмотрены вопросы научно-методического, информационного, технологического и правового обеспечения цифровой трансформации, проектирования и внедрения автоматизированных систем научно-технической информации, библиотечно-информационных систем и технологий, публикационной активности ученых, а также искусственного интеллекта и когнитивных технологий в информатизации.

Материалы конференции будут полезны специалистам в области информатизации и коммуникационных технологий, занимающимся научно-методическим обеспечением информатизации и решением задач построения ИТ-страны, цифровой экономикой, разработкой и внедрением автоматизированных информационных систем управления, систем научно-технической информации, автоматизированных библиотечно-информационных систем и технологий, а также развитием информационной инфраструктуры Беларуси и других стран, реализацией проектов государственных и отраслевых программ в сфере информатизации.

Одобрены программным комитетом и печатаются по решению редакционной коллегии Объединенного института проблем информатики Национальной академии наук Беларуси в виде, представленном авторами.

Научные редакторы:

доктор военных наук, кандидат технических наук, доцент С. В. Кругликов
кандидат технических наук, доцент Р. Б. Григянец
кандидат технических наук, доцент В. Н. Венгерев

БЕЛАРУСКАМОЎНЫ ГАЛАСАВЫ АІ-АСІСТЭНТ

В. В. Дыдо, М. С. Люціч, М. А. Павуціна, А. Я. Драгун,
В. А. Хахлоў, А. С. Трафімаў, Я. С. Зяноўка, Ю. С. Гецэвіч
Аб'яднаны інстытут праблем інфарматыкі НАН Беларусі, Мінск

Прадстаўлена інтэрактыўная платформа «Галасавы АІ-асістэнт», якая ўключае набор пытальна-адказных сістэм на беларускай мове. Іх якаснае функцыянаванне абумоўлена выкарыстаннем высока якасных тэхналогій распазнавання і сінтэзу маўлення новага пакалення, машыннага навучання і дыялогавых сістэм.

Уводзіны

У сучасным асяроддзі адбываецца небывалы рост аб'ёму даных і іх апрацоўка з магчымасцю здабывання каштоўнай інфармацыі дзякуючы разнастайным падзадачам штучнага інтэлекту. Штучны інтэлект можа аналізаваць вялікія аб'ёмы даных у рэальным часе, рабіць прадказанні на іх аснове і выдаваць неабходную інфармацыю, а таксама ўключае ў сябе розныя тэхнікі і метады, дазваляючы камп'ютарным сістэмам ажыццяўляць задачы, якія патрабуюць інтэлектуальных здольнасцяў. Так, пытальна-адказныя сістэмы выкарыстоўваюць метадыкі апрацоўкі натуральнай мовы (*Natural Language Processing, NLP*), распазнавання і сінтэзу маўлення, машыннага навучання, дыялогавыя сістэмы і іншыя алгарытмы для таго, каб разумець пытанні, шукаць адпаведную інфармацыю і генэраваць адказы на аснове наяўных даных. Яны шукаць адпаведную інфармацыю і генэраваць адказы на аснове наяўных даных. Яны імкнуцца вырашыць складаную задачу разумення натуральнай мовы, што з'яўляецца адным з ключавых кампанентаў штучнага інтэлекту. Лабараторыяй распазнавання і сінтэзу маўлення Аб'яднанага інстытута праблем інфарматыкі НАН Беларусі ([URL: http://ssrlab.by/](http://ssrlab.by/)) прадстаўлена інтэрактыўная платформа «АІ-асістэнт» ([URL: https://asistent.by/](https://asistent.by/)), якая змяшчае набор галасавых пытальна-адказных сістэм у выглядзе чат-ботаў у сацыяльнай сетцы Тэлеграм.

1. Архітэктара пытальна-адказных асістэнтаў платформы

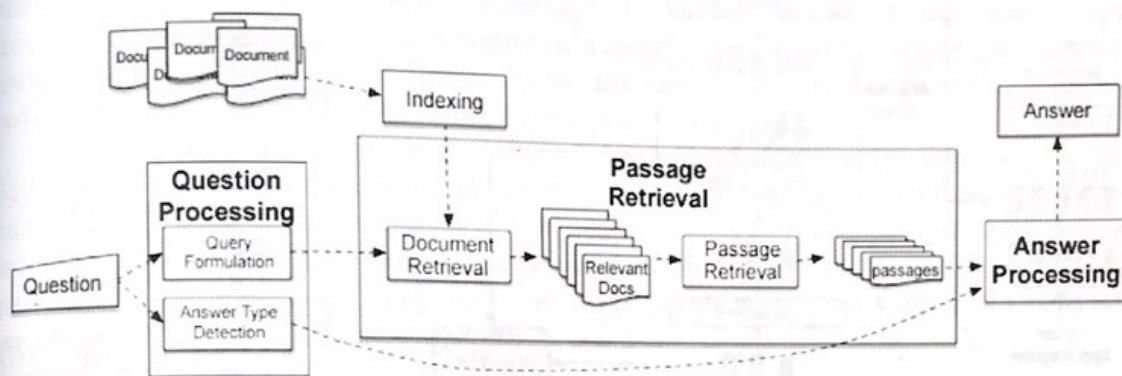
Пытальна-адказныя асістэнты інтэрактыўнай платформы «Галасавы АІ-асістэнт» маюць наступную структуру (мал. 1):

1. Уваходны інтэрфейс. Карыстальнік можа задаць пытанне чат-боту ў вуснай форме ці ў тэкставым фармаце ў тэлеграм-чаце. Галасавыя запыты апрацоўваюцца складанай тэкставых паведамленняў. Таму на дадзеным этапе выкарыстоўваецца беларускамоўная сістэма распазнавання маўлення (БСРМ) для пераўтварэння аўдыясігнала ў тэкставую форму з мэтай падбору найбольш рэлевантнага адказа ў тым выпадку, калі вымаўленае пытанне апрацавана карэктна БСРМ ([URL: https://corpus.by/BelarusianSpeechRecognition/?lang=en](https://corpus.by/BelarusianSpeechRecognition/?lang=en)). Сістэма якасна апрацоўвае гукавыя запыты, што памяншае верагоднасць выдачы адказа, неадпаведнага запыту.

2. Апрацоўка ўваходных даных. Сістэма ажыццяўляе перадапрацоўку пытання, каб зразумець яго сэнс і кантэкст. Гэты крок уключае набор сродкаў апрацоўкі натуральнай мовы і вылучэнне асноўных сэнсавых адзінак у пытанні.

3. Інфармацыйны пошук, ранжыраванне адказаў і іх выманне. На гэтым кроку галасавы асістэнт шукае прыдатную інфармацыю, якая можа ўтрымліваць адказ на пытанне ў анлайн-рэсурсах. Для беларускамоўных пытальна-адказных сістэм выкарысто-

ўваецца моўная мадэль *GPT* (менавіта мадэль *ChatGPT-3,5*), якая з'яўляецца серыяй моўных мадэляў, распрацаваных кампаніяй OpenAI. Мадэль папярэдне навучана на велізарных наборах тэкставых даных. Дзякуючы гэтаму *GPT* можа генэраваць тэкст, які мае сэнс, выкарыстоўвае правільную граматыку і структуру сказаў. Калі ёсць некалькі магчымых адказаў, сістэма можа выкарыстоўваць алгарытм ранжыравання для выбару самага рэлевантнага. *ChatGPT-3,5* падтрымлівае беларускую мову, але якасць адказаў не вельмі добрая. Для яе паляпшэння распрацаваны дадатковы блок машыннага перакладу, у якім усе запыты аўтаматычна перакладаюцца на англійскую мову з дапамогай сістэмы Google Translate. Найбольш дакладны адказ, выдадзены *ChatGPT-3,5*, зноў пыходзіць у блок машыннага перакладу, дзе на выхадзе прадстаўляецца ўжо беларускамоўны варыянт адказу.



Мал. 1. Структура пытальна-адказных асістэнтаў платформы «Галасавы AI-асістэнт»

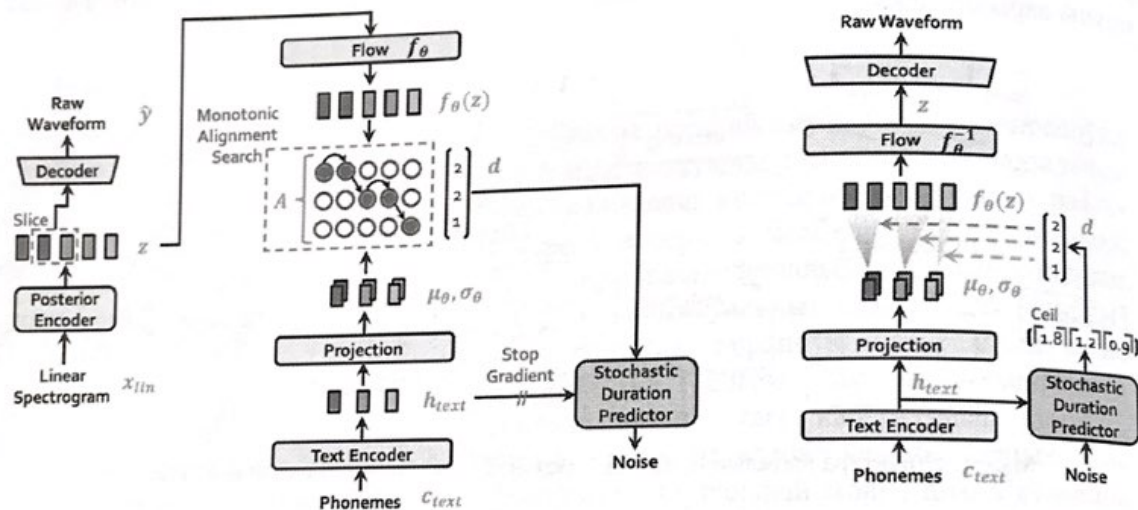
4. Фарміраванне адказу. Галасавы асістэнт генэруе адказ карыстальніку ў выглядзе тэксту, або галасавога паведамлення ў залежнасці ад жадання карыстальніка. Для атрымання галасавога паведамлення выкарыстоўваецца абноўленая мадэль беларускамоўнага анлайн-сінтэзатара маўлення VITS.

2. Беларускаямоўны сінтэзатар маўлення па тэксце новага пакалення

Мадэль VITS (Variational Inference with adversarial learning for Text-to-Speech) уяўляе сабой аднаступеньчатую няўтарэгрэсіённую мадэль пераўтварэння тэксту ў маўленне, здольную генэраваць больш натуральны гук, чым існуючыя двухступеньчатыя мадэлі, такія як Tacotron 2, Transformer TTS ці нават Glow-TTS. Выкарыстоўваючы варыятыўную аснову, VITS мадэлюе латэнтную прастору характарыстык маўлення, адлюстроўваючы ўласцівую зменлівасць і нявызначанасць пры генэраванні маўлення. Наяўнасць спаборнасці навучання ў VITS яшчэ больш удасканальвае працэс сінтэзу. Спаборнае навучанне ўключае ў сабе навучанне сеткі дыскрымінатара для адрознення рэальнай і сінтэзаванай гаворкі, а сетка генератара імкнецца генэраваць маўленне, якое паспяхова падманвае дыскрымінатара.

Такое спаборнае ўзаемадзеянне дапамагае палепшыць агульную якасць і рэалістычнасць сінтэзаваных узораў гаворкі. VITS служыць аўтаномным рашэннем для сінтэзу тэксту ў маўленне, паколькі не патрабуе асобнага вакодэра. Агульная архітэктур VITS адлюстравана на мал. 2. Яна складаецца з кодэра Posterior, кодэра Prior, хітэктур VITS адлюстравана на мал. 2. Яна складаецца з кодэра Posterior, кодэра Prior, дэкодэра Decoder і стахастычнага прадказальніка працягласці. Модулі Posterior Encoder і Decoder Discriminator выкарыстоўваюцца толькі падчас навучання, а не для вываду маўлення. Для Posterior Encoder выкарыстоўваецца 16 рэшткавых блокаў WaveNet, якія складаюцца з слаёў пашыраных скрутак з блокам актывацыі і пропускам сувязі. Задні

энкодэр прымае спектраграмы лагарыфмічнай велічыні ў лінейным маштабе x_{lin} у якасці ўваходных даных і вырабляе латэнтныя зменныя z з 192 каналамі. Ідэя Posterior Encoder заключаецца ў перакладзе аўдыяданых з прасторы mel -спектраграм у прастору нармальнага размеркавання. Менавіта таму ў мадэлі выкарыстоўваецца лінейны пласт па-над Posterior Encoder для атрымання сярэдняй дысперсіі нармальнага апастэрыёрнага размеркавання. Prior Encoder складаецца з Text Encoder, Projection Layer, normalizing Flow і выкарыстоўвае *Monotonic Alignment Search (MSA)*. Як і Posterior Encoder, Prior Encoder накіраваны на адлюстраванне тэкставых даных з прасторы фанем у прастору нармальнага размеркавання.



Мал. 2. Агульная архітэкура VITS

Для распрацоўкі беларускамоўнай сістэмы сінтэзу маўлення па тэксце (БССМТ) істотным крокам з'яўляецца збор даных для навучання мадэлі. Улічваючы абмежаваную даступнасць спецыялізаваных датасэтаў для беларускай мовы, у якасці асноўнай крыніцы даных быў абраны датасэт *CommonVoice*. Ён уяўляе сабой буйную калекцыю галасавых запісаў, сабраных ад добраахвотных удзельнікаў, якія чытаюць сказы на розных мовах. Гэты датасэт даступны для свабоднага выкарыстання і распаўсюджваецца з адкрытай ліцэнзіяй, што робіць яго каштоўнай крыніцай для распрацоўкі лінгваакустычных рэсурсаў на беларускай мове. Недахопам датасэта ад *CommonVoice* з'яўляецца тое, што ён прызначаны для задач распазнання маўлення, а не спецыялізавана для сінтэзу. У сувязі з гэтым запісы ў датасэце могуць быць непрафесійнымі і ўтрымліваць розныя шумы або недахопы, якія негатыўна адбіваюцца на якасці сінтэзу маўлення.

У працэсе збору даных з датасэта *CommonVoice* было праведзена папярэдняе фільтраванне і адбор запісаў, каб пераканацца, што выбраныя даныя адпавядаюць задачы сінтэзу маўлення і добрай якасці. Адабраны аўдыяматэрыял (каля 20 тыс. сказаў) адпавядаў крытэрыю выбару найбольшай колькасці даступнага аўдыя і адносна нізкаму узроўню шуму і артэфактаў. Аналіз выбраных даных ажыццяўляўся з мэтай атрымання статыстычнай інфармацыі і разумення асаблівасцяў беларускай мовы ў кантэксце сінтэзу маўлення. Ён уключаў ацэнку размеркавання фанетычных адзінак, працягласці фраз і іншых характарыстык, якія могуць аказаць уплыў на якасць і натуральнасць сінтэзаванага маўлення.

Мадэль VITS навучалася з дапамогай бібліятэкі *Coqui TTS*, папулярнага набору інструментаў з адкрытым зыходным кодам для БССМТ. *Coqui TTS* прадастаўляе поўны набор інструментаў і ўтыліт для навучання і разгортвання мадэляў TTS. У працэсе навучання мадэль VITS выкарыстала рэгістратар *Weights and Biases*. Гэта платформа для адсочвання і візуалізацыі эксперыментаў машыннага навучання. Яна дазваляе даследчыкам і распрацоўшчыкам рэгістраваць і адсочваць ход навучання, метрыкі і прадукцыйнасць мадэлі ў рэжыме рэальнага часу.

Дзякуючы выкарыстанню магчымасцяў бібліятэкі *Coqui TTS* і інтэграцыі рэгістратара *Weights and Biases* навучанне мадэлі VITS праходзіла з выкарыстаннем надзейнага і трывалага інструментара для распрацоўкі СМТ. Выкарыстанне *Coqui TTS* і рэгістратара *Wandb* спрыяла эфектыўнаму правядзенню эксперыментаў, аптымізацыі мадэлі і маніторынгу прадукцыйнасці на працягу ўсяго працэсу навучання. Для навуметраў адбывалася з дапамогай алгарытма *AdamW* – выпраўленай версіі папулярнага алгарытма аптымізацыі *Adam*. Памер батча для навучання і ацэнкі якасці мадэлі быў роўны 74, час навучання мадэлі склаў 72 гадзіны.

З дапамогай нейронных сетак ажыццяўляліся трэніроўка і навучанне акустычнай базы. Такім чынам, яна створана аўтаматычна і мае вельмі дакладныя вынікі. Аднак дадзены сінтэзатар вялікі па памеры, што можа зніжаць хуткасць падрыхтоўкі тэкставай інфармацыі. Таксама вялікім недахопам лічыцца адсутнасць апрацоўкі лікаў, лічбаў, дат і абрэвіятур, што з'яўляецца аб'ектам распрацоўкі новых метадаў і алгарытмаў па выпраўленні дадзенага недахопа.

3. Функцыянал беларускамоўных пытальяна-адказных асістэнтаў

Інтэрактыўная платформа «Галасавы AI-асістэнт» прадстаўляе сабой прэзентацыйную старонку распрацаваных беларускамоўных пытальяна-адказных асістэнтаў жаночага і мужчынскага полу на трох мовах (беларускай, англійскай, рускай). Кожная пытальяна-адказная сістэма-асістэнт пабудавана з выкарыстаннем тэхналогій распазнавання і сінтэзу маўлення, машыннага перакладу і дыялогавых сістэм. На бягучы момант галасавыя асістэнты даступны ў фармаце тэлеграм-ботаў, дадаткова вядуцца працы па распрацоўцы вэб-версіі.

Перад пачаткам працы карыстальнік можа абраць любога з пяці прапанаваных на сайце асістэнтаў (*AlAlesBot*, *AlAlesiaBot*, *AlAlenaBot*, *AlBorisBot*, *AlKirylBot*), пасля чаго будзе перанакіраваны ў тэлеграм-чат і пачне з ім гутарку. Таксама прадстаўлены два чат-боты жаночага і мужчынскага полу ў прапрацоўцы. Кожная з прапанаваных сістэм мае імя, і ў карыстальніка з'яўляецца ўражанне, што размова вядзецца з рэальным суразмоўцам.

Яшчэ адной адметнай распрацоўкай з'яўляецца чат-бот, які вядзе дыялог згодна прафесійнай накіраванасці (*AsistentBot*). Ён уключае ў сабе трынаццаць роляў: агульны асістэнт, архітэктар, бізнэс-аналітык, фінансавы кансультант, рэкруцёр, праджэкт-мэнаджар, юрыдычны кансультант, маркеталаг, інжынер, праграмоўца, настаўнік і пісьменнік. Сістэма адказвае толькі на беларускай мове. Тэкставы запыт можна даслаць і на іншай мове, але адказ прыйдзе на беларускай, таму што галасавыя запыты распазнаюцца толькі на нашай мове. У пачатку дыялога з *AsistentBot* трэба абраць неабходную спецыяльнасць у прапанаваным спісе ці ўвесці яе з дапамогай спецыялізаваных каманд:

/mode – абярыце памочніка;

/help – паказаць даведачнае паведамленне;

/cancel – адмяніць запыт;

/current_role – даведацца пра выбранага памочніка.

Праверка пытальна-адказных асістэнтаў адбываецца штодзённа: раніцай і ўвечары тэсціроўшчык задае па 1–2 пытанні ботам тэкстам ці голасам. Таксама дадаткова ідзе праверка ўсяго функцыянала бота AsistentBot. Па статыстыцы боты добра адказваюць як на простыя пытанні, напрыклад “Колькі колераў у вясёлцы?”, “Колькі будзе 1+1?”, так і на складаныя: “Распавядзі мне пра самы першы фільм”, “Мне патрэбна гісторыя БНТУ”. Адметнай рысай асістэнтаў з’яўляюцца ветлівасць і прапанова да далейшага дыялогу па тэматыцы, якая зацікавіла карыстальніка. З цягам часу было паляпшана прагаворванне лікаў, гадоў, скарачэнняў, замежных слоў. Зараз гаворка ботаў гучыць лепш, чым было ў пачатку праекта. Праблемы, якія зараз існуюць: перыядычна прыходзіць некарэктны адказ, няправільна прагаворваюць дробныя лікі, некарэктна адказваюць на пытанні з працяжнікамі (напрыклад, “Мінск – гэта”, “Меркурый – гэта”). Кожны тыдзень праца ботаў паляпшаецца.

На сённяшні дзень вядуцца працы па распрацоўцы Web-, iOS- і Android-версій асістэнта для больш зручнага карыстання. Ствараюцца новая версія дызайна афіцыйнай старонкі інтэрактыўнай платформы «Галасавы AI-асістэнт» і лагатып прыстасавання, таксама распрацоўка інтэрфейса на кітайскай і іншых мовах для большай папулярызаванасці чат-ботаў. Немалаважнай задачай з’яўляецца збор водгукаў ад карыстальнікаў пра вынікі дыялогу з сістэмай (ідэі і прапановы прымаем па пошце, адлюстраванай у канцы старонкі платформы). Гэта паспрыяе аналізу памылак або няўдалых адказаў, каб паляпшыць мадэль адказаў ці пашырыць тэхнічныя магчымасці ботаў.

Заклучэнне

Дадзены артыкул апісвае адну з падзадач штучнага інтэлекту – пытальна-адказныя сістэмы на беларускай мове. Мэтай распрацоўкі галасавых асістэнтаў з’яўляецца забеспячэнне эфектыўнага і простага ў выкарыстанні механізма прадастаўлення агульнай інфармацыі і рашэння пытанняў карыстальнікаў на беларускай мове. Інтэрактыўная платформа «Галасавы AI-асістэнт» дазваляе карыстальніку голасам ці ўводам тэксту з клавіятуры задаць пытанне на беларускай мове і атрымаць на яго гукавы ці надрукаваны адказ. За кошт выкарыстання штучнага інтэлекту яна дае магчымасць атрымліваць хуткія, якасныя і дакладныя адказы па розных тэмах. У выніку штодзённага навучання галасавы асістэнт можа як весці гаворку па навуковых тэмах, так і зрабіць забаўляльныя прапановы.

Актуальнасць дадзеных асістэнтаў абумоўлена адсутнасцю канкурэнтаздольных чат-ботаў, якія падтрымліваюць беларускую мову, у той час як для іншых моў існуе велізарная колькасць галасавых асістэнтаў. Распрацоўка прыстасаванняў на беларускай мове робіць іх больш даступнымі для мясцовых карыстальнікаў, прапануючы зручны варыянт пошуку інфармацыі і камунікацыі ў сучаснай глабальнай інтэрнэт-сетцы і выкарыстання камп’ютарных тэхналогій на іх роднай мове.

2. АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ, ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ НАУЧНОЙ, НАУЧНО-ТЕХНИЧЕСКОЙ И ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ

Деев Н. А., Бокуть Л. В., Ковалев М. Я. Интеллектуальный анализ данных и цифровые технологии: важнейшие результаты 2022 года.....	154
Ганчерёнок И. И., Горбачев Н. Н., Хидиров У. Д., Абдихаиров Р. А. Моделирование Международного центра открытых образовательных ресурсов	159
Григянец Р. Б., Успенский Ал. А., Венгеров В. Н. Формирование и ведение единого информационного ресурса по обеспечению инновационной деятельности и трансфера технологий в НАН Беларуси	164
Григянец Р. Б., Венгеров В. Н., Молчан Ж. М., Котов В. И., Успенский А. Ал., Успенский Ал. А., Прибыльский М. С. О повышении компетенций сотрудников НАН Беларуси в сфере трансфера технологий.....	168
Успенский А. Ал., Успенский Ал. А. Новые интернет-инструменты продвижения разработок, продукции и услуг, предоставляемые Республиканским центром трансфера технологий организациям НАН Беларуси.....	172
Горбач Л. А. Медицинские нейронные сети: возможности, ограничения, области применения.....	176
Липницкий С. Ф., Степура Л. В. Поиск и лексико-семантическая обработка научно-технической информации	181
Дравица В. И., Король И. А., Полещук А. В. Цифровые экосистемы идентификации и прослеживаемости товаров в цепях поставок.....	186
Дыдо В. В., Люціч М. С., Павуціна М. А., Драгун А. Я., Хахлоў В. А., Трафімаў А. С., Зяноўка Я. С., Гецэвіч Ю. С. Беларускамоўны галасавы AI-асістэнт	190
Григянец Р. Б., Рабушко К. А. Интеллектуализация информационно-поисковых систем с базами документов неоднородной структуры	195
Слесарова М. М., Латышэвіч Д. І., Танюкевіч Д. П., Назараў У. У., Шаховіч Ю. В., Зяноўка Я. С., Гецэвіч Ю. С. Інтэрактыўныя карты размяшчэння зімавальных ям у рыбалоўных узгоддзях	200
Павлов П. А., Коваленко Н. С. Ресурсно-процессная модель распределенных вычислений при ограниченном числе копий программного ресурса	205