Облікова картка НДДКР

Державний обліковий номер: 0221U100983

Державний реєстраційний номер: 0120U103416

Відкрита

Дата реєстрації: 11-01-2021

1. Етапи виконання

Номер етапу: 1

Назва етапу: Розробити систему базових сервісів семантичного аналізу природномовного опису заданих предметних галузей

Початок етапу: 06-2020

Закінчення етапу: 12-2020

Вид звітного документа: Остаточний звіт

2. Виконавець

Назва організації: Інститут кібернетики імені В.М. Глушкова Національної академії наук України

Код ЄДРПОУ/ІПН: 05417176

Підпорядкованість: Національна академія наук України

Адреса: проспект Академіка Глушкова, 40, м. Київ, Київська обл., 03187, Україна

Телефон: 380445262008

Телефон: 380445264178

E-mail: incyb@incyb.kiev.ua

WWW: http://incyb.kiev.ua/

3. Власник результатів НДДКР (продукції)

Назва організації: Інститут кібернетики імені В. М. Глушкова Національної академії наук України

Код ЄДРПОУ/ІПН: 05417176

Адреса: проспект Академіка Глушкова, буд. 40, м. Київ, Київська обл., 03187, Україна

Підпорядкованість: Національна академія наук України

Телефон: 380445262008

Телефон: 380445264178

E-mail: incyb@incyb.kiev.ua

WWW: http://incyb.kiev.ua/

4. Джерела та напрями фінансування

Підстава для проведення робіт: 34 - договір (замовлення) з центральним органом виконавчої влади, академією наук (головними розпорядниками бюджетних коштів на проведення НДДКР)

КПКВК: 6541030

Напрям фінансування: 2.2 - прикладні дослідження і розробки

Джерела фінансування

Джерело фінансування: 7713 - кошти держбюджету

Фактичний обсяг фінансування за звітний етап: 75 тис. грн.

5. Науково-технічна робота

Назва роботи (укр)

Розробити систему сервісів персоніфікованої інформаційної підтримки наукових досліджень в установах НАН України. Етап 1. Розробити систему базових сервісів семантичного аналізу природномовного опису заданих предметних галузей.

Назва роботи (англ)

Develop a personalized information system to support scientific research in the institutions of NAS of Ukraine. Stage 1. Develop a system of basic services of semantic analysis of natural language description of defined subject fields.

Реферат (укр)

Об`єкт дослідження: методи побудови перспективних знання-орієнтованих комп'ютерних систем підтримки науково-технічної творчості. Мета роботи: розробка сервісів персоніфікованої інформаційної підтримки наукових досліджень в установах НАН України. Створено сервіси: обробки природномовних текстів великого обсягу, зокрема, виділення термінів, контекстів їх використання та фрагментів ієрархії термінів; формування вибірки даних для навчання дистрибутивно-семантичних моделей; валідації метаданих публікацій, підготовлених для розміщення в електронній бібліотеці наукових публікацій України. В межах цих сервісів реалізовано: елементи конвеєру обробки україномовних текстів, зокрема функції та процедури добування текстових даних з електронних документів форматів .docx, .pdf та txt; нормалізацію текстів (лінгвістичний препроцесінг); функцію автоматичного виправлення орфографічних помилок; елементи лексичного аналізу тексту. Сервіс валідації метаданих публікацій може виконувати: збір метаданих в різних схемах, по унікальних ідентифікаторах метаданих setSpec і часових показниках, підтримку базового й розширеного пошуку, використовуючи переходи для полів із всіх архівів, які піддаються індексації. Виконано системно-онтологічний аналіз формування заявок на винаходи на основі онтологічного подання їх семантичного змісту та розроблено відповідні процеси, які описуються онтологічними моделями об’єктів, процесів та задач проблемного простору. Щодо подання знань проектування, пропоновано: слабкі методи рішення задач при відсутності досвіду та точних математичних моделей; сильні методи рішення задач при наявному досвіді проектування аналогічних зразків або точні математичні моделі разом з критеріями якості результатів.

Реферат (англ)

Object of the research: methods of construction of advanced knowledge-based computer systems to support scientific and technical creativity. Purpose of the work: development of services of personalized information support for scientific research in the institutions of NAS of Ukraine. The following services have been created: service for processing large-volume natural language texts, in particular, for highlighting terms, contexts of their usage and fragments of the hierarchy of terms; data sampling service for training distribution-semantic models; service for validation of publication metadata prepared for placement in the electronic library of scientific publications of Ukraine. Within the framework of these services the following options have been implemented: pipeline processing elements of Ukrainian texts, inter alia, the functions and procedures for extracting text data from electronic documents in docx, .pdf and .txt formats; normalization of texts (linguistic preprocessing); the function of automatic spell correction; elements of lexical analysis of the text. The service for validation of publication metadata can perform: collection of metadata in different schemes by unique metadata identifiers and time metrics, support basic and advanced search using field transitions from all archives that can be indexed. The system-ontological analysis of the formation of applications for an invention has been performed based on the ontological representation of their semantic content, as well as the appropriate processes have been developed, which are described by ontological models of the objects, processes and tasks of the problem space. Regarding the presentation of design knowledge, the following are proposed: weak methods of solving problems in the absence of experience and accurate mathematical models; strong methods of solving

problems with existing experience in designing similar samples or accurate mathematical models together with quality criteria for results

Індекс УДК: 004.8 , , 004.8

Коди тематичних рубрик НТІ: 28.23.01 , 28.23.13

6. Науково-технічна продукція (НТП)

НТП 1

Назва продукції (укр): Комплекс засобів інформаційно-технологічної підтримки науково-технічної творчості

Назва продукції (англ): Complex of tools for information and technological support of scientific and technical creativity

Очікувані результати: Методи, теорії, Програмні продукти

Галузь застосування: 72.1 - Дослідження й експериментальні розробки у сфері природничих і технічних наук, установи НАН України

Опис продукції (укр): задля підтримки процесу науково-технічної творчості користувачу надається множина сервісів: • сервіс автоматизованої обробки (виділення термінів, контекстів їх використання та фрагментів ієрархії термінів) природномовних текстів великого обсягу (зокрема, публікацій науковця від 5 до 20 обліково-видавничих аркушів); • сервіс формування вибірки даних; • сервіс валідації метаданих публікацій, підготовлених для розміщення в електронній бібліотеці наукових публікацій України.

Соціально-економічна спрямованість НТП: Підвищення продуктивності праці, Підвищення автоматизації виробничих процесів

Стадія завершеності НТП: Звіт по НДДКР

Впровадження НТП: Демонстраційна версія елементів технології інформаційної підтримки науково-технічної творчості

Строки впровадження: 01.202112.2022

Виробник продукції: ІК НАН України

Споживачі продукції:

Перспективні ринки: Україна

Права інтелектуальної власності: За договорами, Авторські свідоцтва

Форми та умови передачі продукції: Продаж продукції, за договорами

7. Бібліографічний опис

Palagin, O.V., Velychko, V.Yu., Malakhov, K.S., . Shchurov, O.S. Distributional semantic modeling: a revised technique to train term/word vector space models applying the ontology-related approach. Проблеми програмування. 2020. № 2–3, с. 341–351. https://doi.org/10.15407/pp2020.02-03.341

Малахов, К.С. , Щуров О.С. Моделі апаратно-програмного комплексу підтримки наукових досліджень у медицині. Матеріали 24-го міжнародного молодіжного форуму «Радіоелектроніка та молодь в XXI столітті» – Харків, 7–9 квітня 2020 р. том 7+9+10. с.295-297.

Комп'ютерна програма «Когнітивна ІТ платформа поліедр (КІТ ПОЛІЕДР)» ("POLYHEDRON"). Свідоцтво про реєстрацію авторського права на твір. № 96078 Україна. Дата реєстрації 17.02.2020.

Комп'ютерна програма «Трансдисциплінарні онтологічні дослідження операціональних середовищ і процесів» («ТОДОС – ПРОЦЕСИ»). Свідоцтво про реєстрацію авторського права на твір. № 96130. Україна. Дата реєстрації 17.02.2020.

Комп'ютерна програма «Трансдисциплінарна інформаційно-аналітична система ПРИЗМА» ("ТІАС ПРИЗМА"). Свідоцтво про реєстрацію авторського права на твір. № 96125. Україна. Дата реєстрації 17.02.2020.

8. Звітна документація

Кількість сторінок в звіті: 102

Мова звіту: Українська

Кількість файлів у звіті: 1

9. Заключні відомості

Перелік осіб-виконавців

Величко Віталій Юрійович (к.т.н., доц.)

Кургаєв Олександр Пилипович (д.т.н., професор)

Малахов Кирило Сергійович

Новицький Олександр Вадимович (к.т.н.)

Петренко Микола Григорович (д.т.н., с.н.с.)

Стрижак Олександр Євгенович (д. т. н.)

Керівник організації:

Сергієнко Іван Васильович (д. ф.-м. н., акад.)

Керівники роботи:

Палагін Олександр Васильович (д. т. н., професор)

Керівник відділу реєстрації наукової діяльності УкрІНТЕІ

Юрченко Т.А.