Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0824U002771

Особливі позначки: відкрита

Дата реєстрації: 26-07-2024

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Смиш Олег Русланович
2. Oleh R. Smysh

Кваліфікація: 122

Ідентифікатор ORCID ID: 0000-0002-8074-9745

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 122

Назва наукової спеціальності: Комп'ютерні науки

Галузь / галузі знань: інформаційні технології

Освітньо-наукова програма зі спеціальності: Комп'ютерні науки

Дата захисту: 03-10-2024

Спеціальність за освітою: Інженерія програмного забезпечення

Місце роботи здобувача:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): ДФ 26.314.003

Повне найменування юридичної особи: Інститут програмних систем Національної академії наук України

Код за ЄДРПОУ: 05540149

Місцезнаходження: проспект Академіка Глушкова, буд. 40, корп. 5, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Інститут програмних систем Національної академії наук України

Код за ЄДРПОУ: 05540149

Місцезнаходження: проспект Академіка Глушкова, буд. 40, корп. 5, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 20.54

Тема дисертації:

1. Створення рекомендаційної системи для розв’язування математичних задач, записаних природною українською мовою

2. Creating a recommendation system for solving mathematical problems that are written in Ukrainian language

Реферат:

1. Об’єктом дослідження є оброблення текстових даних українською мовою, охоплюючи створення та використання баз знань для логічного виведення висновків та аналізу несуперечності даних за допомогою онтології. Метою роботи є заповнення прогалин, що є в доступних на сьогодні ресурсах для самонавчання, за допомогою створення рекомендаційної системи, зокрема в контексті перевірки точності та коректності математичних задач, уможливлення роботи із задачами, що подано українською мовою та автоматизацією процесу формування кроків розв’язку задач. Методи наукового дослідження, що використано в роботі: емпіричний та евристичний методи. Також використано методи математичної логіки, математичного аналізу, дискретної математики, функційного програмування. Додатково: методи системного підходу, зокрема аналіз, синтез (структурний, параметричний) та декомпозиція. У науковій роботі використано також

спеціальні методи. А саме: токенізація, лематизація, розмічування частин мови, видобування інформації. У процесі проведених досліджень, отримано такі наукові результати: • уперше розроблено метод видобування значущої інформації з текстів україномовних планіметричних задач, що сформовано на основі проведеного кількісного аналізу зібраних планіметричних задач та за допомогою інструментів обробки природної української мови, для реалізації автоматичного розв'язування україномовних планіметричних задач; • уперше запропоновано метод автоматичної перевірки на правильність та несуперечність даних, які отримано під час розв'язування задачі та даних, що подано користувачем, формуючи внутрішнє представлення розв'язку задачі як побудову терма в теорії типів за допомогою функційної мови програмування Lean. • уперше побудовано метод автоматичного генерування кроків для розв'язування математичних задач, який розроблено на основі логічного виведення висновків, що отримано від використання онтологій з інструментами логічного виводу. Створено рекомендаційну систему, що уможливлює підвищення ефективності навчального процесу для учнів, абітурієнтів, студентів, вчителів і викладачів. Фактично розроблено систему, що здатна автоматизовано проаналізувати умову математичних задач, поданих українською мовою, та генерувати кроки їхнього розв'язку у формі, що є зрозумілою. Це уможливлює спростити та унаочнити процес навчання та допомогти уникнути типових помилок під час розв'язування задач. Запропоновано методи та підходи роботи з українським текстом планіметричних задач, які базуються на використанні засобів обробки природної української мови та на проведеному кількісному аналізі текстів планіметричних задач. Реалізовано метод, який аналізує текст задачі та використовує предметну планіметричну онтологію. Під час апробації результатів дослідження показано, що середній результат розв'язання 50 планіметричних задач великими мовними моделями становить 53,67 %, тоді як запропонована рекомендаційна система розв'язує всі ці 50 задач. Тому запропоновані методи для вичленення інформації з тексту та застосування планіметричної онтології надають майже вдвічі краще справляється з розв'язанням планіметричних задач, ніж доступні на сьогодні великі мовні моделі. Також запропоновано метод виправлення помилок лематизації в українській мові за допомогою використання словника, що сприяло підвищенню фінальної точності лематизації на 1,66 %. Застосування онтологій та методів логічного виведення уможливило створеній системі перевіряти отримані розв'язки на несуперечність, що є критичним у точних науках. Впровадження методу з використанням функційної мови програмування для цих цілей надало коректність даних. Отже, результати дослідження мають практичне застосування та здатні сприяти покращенню освітніх стандартів в Україні, збільшенню кількості високоякісних рішень для математики, а також наданню інструмента для підтримування користувачів у процесі навчання. Науково-практичні результати дослідження впроваджено в навчальну діяльність факультету інформатики Національного університету «Києво-Могилянська академія», зокрема на кафедрі мультимедійних систем. Надалі, результати дисертаційної роботи також планується впровадити під час вивчення інших дисциплін, аби сприяти поглибленню знань та аналізу автоматизованих систем за допомогою використання запропонованої рекомендаційної системи, що здатна працювати з україномовними текстами математичних задач, розв'язувати їх самостійно та допомагати з розв'язуванням кінцевим користувачам.

2. The object of the study is text data processing in Ukrainian, including the creation and use of knowledge bases for logical reasoning and analysis of data consistency using ontology. The purpose of the study is to fill the gaps in the currently available self-study resources by creating a recommendation system, in particular in the context of checking the accuracy and correctness of mathematical problems, enabling work with problems presented in Ukrainian, and automating the process of forming steps for solving problems. Research methods used in the study: empirical and heuristic methods. Methods of mathematical logic, mathematical analysis, discrete mathematics, and functional programming. Additionally: methods of the systematic approach, including analysis, synthesis (structural, parametric) and decomposition. Special methods are also used in the research work: tokenization, lemmatization, part-of-speech markup, and information extraction. The scientific novelty of the results obtained in the thesis is as follows: • for the first time, a method for extracting meaningful information from the texts of Ukrainian-language planimetry problems was developed, based on the frequency analysis of the collected

planimetry problems and using natural language processing tools for Ukrainian to implement an automatic solving of Ukrainian-language planimetry problems; • for the first time, a method of automatic verification of the correctness and consistency of the data that is obtained during the solving process of the problem and the data provided by the user, forming an internal representation of the problem's solution as the construction of a term of type theory using the Lean functional programming language. • for the first time, a method for automatic generation of steps for solving mathematical problems was built, that is based on logical inference obtained from the use of ontologies with the logical inference tools. A recommendation system has been created that makes it possible to improve the efficiency of the educational process for pupils, applicants, students, teachers, and lecturers. In fact, a system has been developed that can automatically analyze the conditions of mathematical problems presented in Ukrainian and generate steps for solving them in a form that is understandable. This makes it possible to simplify and visualize the learning process and help avoid common mistakes when solving problems. A method has been implemented that analyzes the text of the problem and uses the planimetry ontology. During the testing of the study results, it was shown that the average result of solving 50 planimetry problems by using large language models is 53.67 %, while the proposed recommendation system solves all 50 of these problems. Therefore, the proposed methods for extracting information from the text and applying the planimetry ontology provide almost twice as better solution to planimetry problems as currently available large language models. There is also a method for correcting lemmatization errors in the Ukrainian language by using a dictionary, which helped to increase the final lemmatization accuracy by 1.66 %. The use of ontologies and methods of logical inference allowed the system to check the obtained solutions for consistency, which is critical in the exact sciences. The implementation of a method using a functional programming language for these purposes ensured the correctness of the data. Thus, the results of the study have practical application and can contribute to improve educational standards in Ukraine, increasing the number of high-quality mathematics solutions, and providing a tool to support users in the learning process. The scientific and practical results of the research have been implemented in the educational activities of the Faculty of Informatics of the National University of Kyiv-Mohyla Academy, in particular at the Department of Multimedia Systems. In the future, the results of the dissertation are also planned to be implemented in the study of other disciplines to help deepen knowledge and analysis of automated systems by using the proposed recommendation system that can work with Ukrainian-language texts of mathematical problems, solve them independently and help end users to solve them.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- Жежерун О. П., Смиш О. Р. Автоматизація розв'язування задач з планіметрії, записаних природною українською мовою. Проблеми програмування. 2020. № 4. С. 71–80. // Категорія наукового фахового видання Б. URL: https://doi.org/10.15407/pp2020.04.071
- Димченко О. В., Смиш О. Р., Жежерун О. П. Графічний інтерфейс для рекомендаційної системи. Наукові записки НаУКМА. Комп'ютерні науки. 2021. Т. 4. С. 93–97. // Категорія наукового фахового видання Б. URL: https://doi.org/10.18523/2617-3808.2021.4.93-97
- Жежерун О. П., Смиш О. Р., Пруднікова А. О. Підходи до побудови виводу в онтологічній базі знань. Наукові записки НаУКМА. Комп'ютерні науки. 2023. Т. 6. С. 17–23. // Категорія наукового фахового видання Б. URL: https://doi.org/10.18523/2617-3808.2023.6.17-23
- Смиш О. Р. Організація інтерфейсу для взаємодії з користувачем в рекомендаційній системі. Наука і техніка сьогодні. Серія «Техніка». 2024. № 3 (31). С. 980–989. // Категорія наукового фахового видання

Б. URL: https://doi.org/10.52058/2786-6025-2024-3(31)-980-989

- Смиш О., Загорулько А. Використання мови Lean для перевірки на несуперечність математичних задач у навчальній рекомендаційній системі. Вимірювальна та обчислювальна техніка в технологічних процесах. 2024. № 1. С. 32–38. // Категорія наукового фахового видання Б. URL: https://doi.org/10.31891/2219-9365-2024-77-4
- Смиш О. Р., Жежерун О. П., Резніченко В. А. Створення рекомендаційної системи для розв'язування математичних задач, записаних природною українською мовою. UkrProgAsp 2022-1: 1st Conference of Young Scientists in Programming = УкрПрогАсп-2022-1: 1-а конференція молодих вчених з програмування (26.10.2022, Київ) : збірник матеріалів. Київ : Інститут програмних систем Національної академії наук України, 2022. С. 19–20
- Смиш О. Р., Жежерун О. П. Виправлення помилок лематизації тексту за допомогою словника. «Наукова весна» 2023 : матеріали XIII Міжнародної науково-технічної конференції студентів, аспірантів та молодих вчених (Дніпро, 1–3 березня 2023 року). НТУ «ДП», 2023. С. 193–194
- Смиш О. Р., Жежерун О. П., Резніченко В. А. Формування планіметричної онтології для системи автоматичного розв'язування задач, що записані природною українською мовою. UkrProgAsp 2023-2: 2nd Conference of Young Scientists in Programming = УкрПрогАсп-2023-2: 2-га конференція молодих вчених з програмування (17.10.2023, Київ) : збірник матеріалів. Київ : Інститут програмних систем Національної академії наук України, 2023. С. 32
- Смиш О. Р. Система для розв'язування задач з геометрії : текстова частина магістерської роботи за спеціальністю 121 «Інженерія програмного забезпечення» / Міністерство освіти і науки України, Національний університет «Києво-Могилянська академія», Кафедра мережних технологій факультету інформатики. Київ, 2020. 41 с. URL: https://ekmair.ukma.edu.ua/handle/123456789/18282

Наукова (науково-технічна) продукція: методи, теорії, гіпотези; програмні продукти, програмно-технологічна документація

Соціально-економічна спрямованість: забезпечення промисловості чи населення новим видом інформаційно-комунікаційних послуг

Охоронні документи на ОПІВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Жежерун Олександр Петрович

2. Oleksandr P. Zhezherun

Кваліфікація: к. ф.-м. н., доц., 01.05.03

Ідентифікатор ORCID ID: 0000-0002-4034-6730

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Києво-Могилянська академія"

Код за ЄДРПОУ: 16459396

Місцезнаходження: вул. Г. Сковороди, буд. 2, Київ, 04070, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Резніченко Валерій Анатолійович
2. Valerii A. Reznichenko

Кваліфікація: к. ф.-м. н., с.н.с., 01.05.01

Ідентифікатор ORCID ID: 0000-0002-4451-8931

Додаткова інформація:

Повне найменування юридичної особи: Інститут програмних систем Національної академії наук України

Код за ЄДРПОУ: 05540149

Місцезнаходження: проспект Академіка Глушкова, буд. 40, корп. 5, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Марченко Олександр Олександрович
2. Oleksandr O. Marchenko

Кваліфікація: д. ф.-м. н., професор, 01.05.01

Ідентифікатор ORCID ID: 0000-0002-5408-5279

Додаткова інформація:

Повне найменування юридичної особи: Київський національний університет імені Тараса Шевченка

Код за ЄДРПОУ: 02070944

Місцезнаходження: вул. Володимирська, буд. 60, Київ, 01033, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Пасічник Володимир Володимирович

2. Volodymyr V. Pasichnyk

Кваліфікація: д.т.н., професор, 01.05.01

Ідентифікатор ORCID ID: 0000-0002-5231-6395

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, буд. 12, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Рогушина Юлія Віталіївна

2. Yuliya V. Rohushyna

Кваліфікація: к. ф.-м. н., доц., 01.05.03

Ідентифікатор ORCID ID: 0000-0001-7958-2557

Додаткова інформація:

Повне найменування юридичної особи: Інститут програмних систем Національної академії наук України

Код за ЄДРПОУ: 05540149

Місцезнаходження: проспект Академіка Глушкова, буд. 40, корп. 5, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Новицький Олександр Вадимович

2. Oleksandr V. Novytskyi

Кваліфікація: к.т.н., 05.13.06

Ідентифікатор ORCID ID: 0000-0002-9955-7882

Додаткова інформація:

Повне найменування юридичної особи: Інститут програмних систем Національної академії наук України

Код за ЄДРПОУ: 05540149

Місцезнаходження: проспект Академіка Глушкова, буд. 40, корп. 5, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради	Шевченко Віктор Леонідович
Власне Прізвище Ім'я По-батькові головуючого на засіданні	Шевченко Віктор Леонідович
Відповідальний за підготовку облікових документів	Янченко Олена Станіславівна
Реєстратор	УкрІНТЕІ
Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності	Юрченко Тетяна Анатоліївна