Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0420U102047

Особливі позначки: відкрита

Дата реєстрації: 24-11-2020

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Ковилін Єгор Романович

2. Kovilin Yehor

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 01.05.02

Назва наукової спеціальності: Математичне моделювання та обчислювальні методи

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 11-11-2020

Спеціальність за освітою: Інформаційні управляючі системи та технології

Місце роботи здобувача: Вищий навчальний заклад "Університет імені Альфреда Нобеля"

Код за ЄДРПОУ: 20201672

Місцезнаходження: вул. Січеславська Набережна, 18, м. Дніпро, Дніпропетровський р-н., Дніпропетровська обл., 49000, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): Д 08.084.01

Повне найменування юридичної особи: Національна металургійна академія України

Код за ЄДРПОУ: 02070766

Місцезнаходження: проспект Гагаріна, 4, м. Дніпро, Дніпропетровський р-н., Дніпропетровська обл., 49600, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

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

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Гагаріна, 72, м. Дніпро, Дніпропетровський р-н., Дніпропетровська обл., 49010, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 28.29.25, 28.29.67.11, 28.17.19

Тема дисертації:

1. Модель генерації відповідей в пошукових системах на основі неструктурованої бази знань

2. Model of answers generating in the search engines based on an unstructured knowledge base

Реферат/анотація:

1. Дисертаційну роботу присвячено вирішенню актуальної науково-прикладної задачі розробки моделі для автоматизації обробки семантично-неструктурованих документів для генерації відповідей в пошукових системах. На основі розроблених математичних моделей автоматичного отримання семантичних характеристик тексту і автоматичної класифікації вхідних даних за ступенем їх семантичної зв'язності, запропоновано модель і алгоритм генерації відповідей на запит користувача на основі неструктурованої бази знань. У вигляді відкритого програмного API реалізовані: семантичнаа модель тексту, модель автоматичної класифікації тестів за ступенем їх зв'язності та модель автоматичної генерації текстів на основі неструктурованої бази знань. Використання отриманих у роботі результатів дозволяє отримувати семантичні моделі текстів без залучення ручної семантичної розмітки або лінгвістичних знань і спростити роботу

користувача із пошуковими системами. Ключові слова: математична модель пошукової системи, семантична модель, неструктуровані текстові знання, генерація відповідей, комп'ютерні моделі представлення знань.

2. The dissertation is devoted to the solution of the actual scientific and applied problem of model development for automation of processing the semantically unstructured documents for the answers generation in the search engines. The dissertation analyzes the existing approaches to the construction of applied models of the texts in natural language. Existing approaches require the involvement of significant volumes of linguistic knowledge, preliminary construction of ontological markup or manual creation of the semantic dictionaries of knowledge, which significantly reduces the adaptability of models and complicates their applied use. The existing software systems rely in their work on the extraction of knowledge using ontological bases, which makes it impossible to use them for processing the unstructured knowledge bases. Based on the results of the analysis, it became obvious that there is a need to develop the models that allows the automating process of obtaining and presenting semantic models of specialized texts without the need to involve linguistic knowledge or the formation of the preliminary semantic markups or dictionaries. The first developed model was a mathematical model for obtaining the semantic characteristics of a specialized text. The created model is innovative, and uses in its work the method of latent semantic analysis, clustering and spatial analysis. The created model makes it possible to obtain a semantic model of a text without using a previously created semantic structure or semantic dictionaries, which made it possible to fully automate the process of obtaining quantitative values of the semantic characteristics of the text. The carried out tests showed that, despite the use of the frequency characteristics of the text, the created model depends precisely on the semantics of a document, and not on its frequency portrait. In addition, the created model makes it possible to reliably combine semantically related terms into semantic text labels, while establishing a semantic connection with a set of sentences from the source text, which is a prerequisite for building an automatic text generation process. The second developed model was the model of automatic classification of incoming documents, which uses the quantitative characteristics from the semantic model of text in the process of its work. The created model allows system to filter the incoming documents, thus protecting the created knowledge base from being filled with inappropriate texts. The carried out tests showed that the accuracy of the system, built on the basis of the created model, is 90%, which is sufficient for the correct operation of this stage. The last model created is the unstructured knowledge base response generation model. This model works on the basis of the first and second models and allows to generate the new text knowledge relevant to the user's request. The tests have shown that the use of the developed model allows improvements in the semantic quality of the set of candidate texts for generation by 1.7 times, in comparison with the direct solution of the search problem. The expert assessments carried out by the ball method showed a value of 0.839, which proves the applicability and adequacy of the created model. The developed software application was implemented in the city municipal cultural institution «Centralized system of libraries for children» in Dnipro as a search tool for electronic text processing, in LLC «Sital Ukraine» as a tool of automatic generation of text instructions and in JSC «DniproAzot» as a tool to improve search processes. The results of the dissertation are published in 14 scientific papers, including 8 articles - in journals recommended by the Ministry of Education and Science of Ukraine for publication of dissertations and foreign publications, and 6 - in abstracts and papers of international and national conferences.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки:

Стратегічний пріоритетний напрям інноваційної діяльності:

Підсумки дослідження:

Публікації:

Наукова (науково-технічна) продукція:

Соціально-економічна спрямованість:

Охоронні документи на ОПІВ:

Впровадження результатів дисертації:

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Волковський Олег Степанович

2. Volkovskij Oleg

Кваліфікація: к.т.н., 05.14.06, 05.16.01

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Кіріченко Людмила Олегівна

2. Kirichenko Lyudmyla Olegivna

Кваліфікація: д. т. н., 01.05.02

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Власне Прізвище Ім'я По-батькові:

1. Шинкаренко Віктор Іванович

2. Shynkarenko Viktor Ivanovych

Кваліфікація: д. т. н., 01.05.03

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Рецензенти

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Гнатушенко Вікторія Володимирівна

Власне Прізвище Ім'я По-батькові головуючого на засіданні Гнатушенко Вікторія Володимирівна

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності Юрченко Т.А.