Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0415U003874

Особливі позначки: відкрита

Дата реєстрації: 07-07-2015

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Діхтяренко Олександр Васильович
2. Dikhtiarenko Oleksandr Vasylovych

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 05.13.06

Назва наукової спеціальності: Інформаційні технології

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 30-06-2015

Спеціальність за освітою: 8.092104

Місце роботи здобувача: Київський національний університет будівництва та архітектури

Код за ЄДРПОУ: 02070909

Місцезнаходження: 03680, м. Київ, Повітрофлотський проспект, 31

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 26.056.01

Повне найменування юридичної особи: Київський національний університет будівництва і архітектури

Код за ЄДРПОУ: 02070909

Місцезнаходження: просп. Повітрофлотський, 31, м. Київ, Київська обл., 03680, 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: Київський національний університет будівництва та архітектури

Код за ЄДРПОУ: 02070909

Місцезнаходження: 03680, м. Київ, Повітрофлотський проспект, 31

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 28.23.27

Тема дисертації:

1. Інформаційна технологія визначення нечітких збігів в контенті електронних документів
2. Information technology of matching approximate duplicates within the content of electronic documents

Реферат/анотація:

1. Дисертація присвячена проблемі плагіату наукових та інших робіт, яка завдяки росту комп'ютерної грамотності населення та проникненню Інтернету у всі сфери життя набуває все більшого масштабу. В рамках дисертаційного дослідження розроблені моделі та методи, які дозволяють визначати нечіткі збіги в тексті, зображеннях та таблицях документу, навіть при значних структурних змінах контенту. Запропоновано концептуальну модель технології визначення нечітких збігів, модель індексу документу та модель нечіткого збігу. Для підготовки тексту запропоновано використовувати засоби виправлення помилок в словах, відкидання стоп-слів і стоп-фраз, методи канонізації, десинонімізації та методи заміни антонімів. Розроблено метод створення індексу документу з використанням локально-чутливого хешування, та метод фільтрації помилкових збігів. Для графічних зображень запропоновано методи фрагментації та визначення базового кута повороту зображення. Для таблиць запропоновано методи визначення шапки таблиці та

індексації по стовпцях і рядках. Удосконалено метод кластеризації документів з використанням частоти вживання слова в документі як ознаки кластеризації. Запропоновано варіанти створення моделей розширеного індексу документу, використання яких дозволить пришвидшити виконання процесу пошуку. Розроблена архітектура системи визначення нечітких збігів як програмного комплексу та запропоновано варіанти реалізації компонентів системи. Реалізована система у вигляді програмного забезпечення дозволяє виявляти запозичення у будь-яких типах документів. Система може бути впроваджена у ВНЗ, що дозволить підвищити рівень підготовки фахівців. У наукових періодичних виданнях система дозволить запобігти випадкам шахрайства і привласнення чужих робіт.

2. The thesis is devoted to the problem of plagiarism scientific and other works, which is growing in scale due to the growth of computer literacy and Internet penetration in all spheres of life. Within the framework of the research, new models and methods were developed. These models and methods allow to find fuzzy matches in the text, images and tables within the document, even if content structure was modified. A conceptual model of technology definition of fuzzy matches, the model index of the document and the model of fuzzy matches was developed. For the texts preparation proposed an approach for correcting errors in words, discarding stop-words and stop-phrases, methods of canonization and replacement techniques for synonyms and antonyms. Developed a method for creating document indexes using locality-sensitive hashing, and method for filtering spurious matches. For graphic images, the methods of fragmentation and the definition of the reference rotation angle of the image are proposed. Developed a method of determining the header of the table and a way to index the table by columns and rows. Also a method of clustering documents using the frequency of use of words in the document as signs of clustering has been enhanced. Offered the techniques of creating models of extended index document, which will speed up the execution of the search process. Developed the system architecture definition of fuzzy matches. Implementation of the system as the software allows to identify matches in all types of documents. The system can be used by higher educational institutions that will improve the level of training of specialists. It can be used by scientific periodicals that will help to prevent fraud and misappropriation of other people's work as well as in other structures.

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. Білощицький Андрій Олександрович
2. Biloshchytskyi Andrii Oleksandrovych

Кваліфікація: д.т.н., 05.13.22

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. Тесля Юрій Миколайович
2. Тесля Юрій Миколайович

Кваліфікація: д.т.н., 05.13.06

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. Бадьоріна Любов Миколаївна
2. Бадьоріна Любов Миколаївна

Кваліфікація: к.т.н., 05.13.06

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	Міхайленко Віктор Мефодійович
Власне Прізвище Ім'я По-батькові головуючого на засіданні	Міхайленко Віктор Мефодійович
Відповідальний за підготовку облікових документів	
Реєстратор	
Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності	Юрченко Т.А.

