Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0412U002876

Особливі позначки: відкрита

Дата реєстрації: 13-06-2012

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Лан Чуньлінь

2. Lang Chunlin

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 05.13.05

Назва наукової спеціальності: Комп'ютерні системи та компоненти

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 11-06-2012

Спеціальність за освітою: 8.05010201

Місце роботи здобувача: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: 03056, м.Київ, пр.Перемоги, 37

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 26.002.02

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського" Інститут енергозбереження та енергоменеджменту

Код за ЄДРПОУ: 247571500

Місцезнаходження: вул. Борщагівська 115, м. Київ, Київська обл., 03056, 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: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: 03056, м.Київ, пр.Перемоги, 37

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 50.41.25

Тема дисертації:

1. Методології та програмні засоби класифікації природомовних текстових документів

2. Methodology and software classify natural language text documents

Реферат/анотація:

1. Дисертаційна робота присвячена вирішенню задачі автоматичної ідентифікації мови і класифікації текстових природомовних документів. Запропоновано метод автоматичної ідентифікації мови з використанням статистик N-грам, проведено порівняльний аналіз різних способів класифікації текстових документів з метою вибору оптимальних за точністю і повнотою, запропонована класифікація природомовних текстових документів з використанням розробленого методу статистик N-грам, запропоновано метод автоматичного створення багатотемної класифікації текстових документів в реальному часі, створено програмний модуль для ідентифікації мови та класифікації багатотемних текстових документів. Запропонований спосіб класифікації текстів і документів дозволив підвищити точність і швидкість класифікації, розробити відповідні програмні засоби для використання в системах автоматичної обробки текстів в багатомовних інформаційних системах.

2. The thesis is devoted to solving the problem of automatic language identification and classification of natural language text documents. The method for automatic identification of languages using statistical N-grams, comparative analysis of different methods of classification of text documents in order to choose optimal precision and recall, the proposed classification natural language text documents using the method developed by statistical N-grams, the method automatically classify text documents in real time, created a software module for the identification and classification natural language text documents. The proposed method of classification of text documents allowed to improve accuracy and speed of classification, to develop appropriate software for use in automatic processing of texts in multilingual information systems. Keywords: automatic language identification, classification of text documents, natural language, N-grams, multi-label classification.

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. Зайцев Володимир Григорович

2. Zajschev Volodimir Grigirjevich

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

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Бузовський Олег Володимирович

2. Бузовський Олег Володимирович

Кваліфікація: д.т.н., 05.13.05

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. Селігей Олександр Минович

2. Селігей Олександр Минович

Кваліфікація: к.т.н., 05.13.05

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 Луцький Георгій Михайлович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Луцький Георгій Михайлович

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.