

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ КІБЕРБЕЗПЕКИ, КОМП'ЮТЕРНОЇ
ТА ПРОГРАМНОЇ ІНЖЕНЕРІЇ



МІЖНАРОДНА
НАУКОВО-ТЕХНІЧНА КОНФЕРЕНЦІЯ
«ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ
ЛІНГВІСТИЧНОГО АНАЛІЗУ»

19-20 жовтня 2021 року

Тези доповідей

Київ 2021

Міжнародна науково-технічна конференція «Інтелектуальні технології лінгвістичного аналізу»: Тези доповідей. – К.: НАУ, 2021. – 70 с.

Збірник містить тези доповідей, які були представлені на конференції «Інтелектуальні технології лінгвістичного аналізу».

В доповідях розглянуто дослідження методів, алгоритмів та комп'ютерних технологій лінгвістичного аналізу, принципів побудови інтелектуальних мовно-інформаційних та пошукових систем, методів формалізації та обробки знань, обмін досвідом їх застосування в науці та освіті. Для фахівців з комп'ютерної лінгвістики.

Редакційна колегія:

Литвиненко О.Є. – д.т.н., професор, завідувач кафедри комп'ютеризованих систем управління НАУ (Україна, Київ)

Ланде Д.В. – д.т.н., завідуючий відділом Інституту проблем реєстрації інформації НАН України (Україна, Київ)

Додонов О.Г. – д.т.н., професор, заступник директора Інституту проблем реєстрації інформації НАН України (Україна, Київ)

Затверджено до друку науково-технічною радою Факультету кібербезпеки, комп'ютерної та програмної інженерії Національного авіаційного університету (протокол №9 від 11.10.2021 р.)

© Національний авіаційний університет, 2021

ЗМІСТ

Ланде Д.В., Дмитренко О.О. ВИКОРИСТАННЯ НАПРАВЛЕНИХ ЗВАЖЕНИХ МЕРЕЖ ТЕРМІНІВ ДЛЯ ВИЗНАЧЕННЯ СТУПЕНЯ ПОДІБНОСТІ ТЕКСТІВ	7
Чернецький Д.В., Конрад Т.І. КОМП'ЮТЕРНИЙ ЗІР. СУЧАСНИЙ СТАН ТА ЗАСТОСУВАННЯ .8 Чут Є.В., Ребіцький М.В., Конрад Т.І. ОГЛЯД МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ	9
Івашенко В.В., Конрад Т.І. ЗАСТОСУВАННЯ ОНТОЛОГІЙ ДЛЯ ОБРОБКИ ПРИРОДНОЇ МОВИ З МЕТОЮ УПРАВЛІННЯ ВИМОГАМИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ.....	10
Дорожинська А.В. ОЦИФРУВАННЯ ТЕРМІНОЛОГІЧНОГО СЛОВНИКА	11
Dehtiar Yu.V., Dehtiar N.V. IMPROVEMENTS TO DISPATCH OF MATERIAL FLOWS OF TECHNOLOGICAL UNITS	12
Тенягін Д.Д. ШТУЧНИЙ ІНТЕЛЕКТ У СУЧАСНОМУ БІЗНЕСІ	14
Гурбанов Т.А. ТЕХНОЛОГІЇ ПОРІВНЯЛЬНОГО АНАЛІЗУ ЕЛЕКТРОННИХ ТЕКСТІВ ЯК ЗАСІБ БОРОТЬБИ З ПЛАГІАТОМ.....	15
Луцик О.Р., Срібна В.Д. СПАМ-ФІЛЬТР НА ОСНОВІ ТЕОРЕМИ БАЕСА	17
Стенякін І.А. СТИЛОМЕТРИЧНІ ЗАСОБИ АУТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ В ОНЛАЙНОВИХ ЕЛЕКТРОННИХ СИСТЕМАХ НАВЧАННЯ.....	18
Вавіленкова А.І. ПОШУКОВІ АЛГОРИТМИ ЯК ОСНОВА РОБОТИ ІНФОРМАЦІЙНО-ПОШУКОВИХ СИСТЕМ.....	19
Зубок В.Ю., Савельєв Д.В. РИЗИКИ РОЗПОДІЛЕНОЇ ОБРОБКИ ІНФОРМАЦІЇ В ГЛОБАЛЬНИХ ІНФОРМАЦІЙНО-КОМП'ЮТЕРНИХ СИСТЕМАХ.....	20
Тачиніна О.М., Лисенко О.І. МЕТОДИКА НАЛАШТУВАННЯ ПАРАМЕТРІВ ЦИФРОВОЇ СИСТЕМИ СТАБІЛІЗАЦІЇ БЕЗПЛОТНОГО ЛІТАЛЬНОГО АПАРАТУ	22

УДК 004.912

Ланде Д.В., д.т.н., Дмитренко О.О.

*Інститут проблем реєстрації інформації НАН України
Національний технічний університет «Київський політехнічний
інститут імені Ігоря Сікорського»*

ВИКОРИСТАННЯ НАПРАВЛЕНИХ ЗВАЖЕНИХ МЕРЕЖ ТЕРМІНІВ ДЛЯ ВИЗНАЧЕННЯ СТУПЕНЯ ПОДІБНОСТІ ТЕКСТІВ

Ера стрімкого розвитку інформаційно-телекомунікаційних технологій та інформаційного простору характеризується сплеском неструктурованих та дублюючих даних, зокрема текстових, що значно ускладнює пошук необхідної інформації. Тому актуальною є задача формалізації текстових даних й представлення їх у формі, яка була б зручною для комп'ютеризованої обробки з метою виявлення подібної або дублюючої інформації.

Мета роботи – представити методику визначення ступеня подібності між текстовими документами, що базується на використанні направлених зважених мереж термінів [1], де вузлами таких мереж є ключові терміни тексту, а ребра – семантико-семантичні зв'язки між цими термінами у тексті.

Побудова мереж термінів здійснюється в декілька етапів [1], що включають попередню обробку текстових даних, виокремлення ключових термінів, застосування алгоритму графа горизонтальної видимості для встановлення ненаправлених зв'язків між термінами, а також подальше встановлення напрямків зв'язків та їх вагових значень. Подальше порівняння за допомогою міри Фробеніуса матриць направлених зважених мереж, що отримані для різних текстів, дозволяє визначити семантичну близькість відповідних текстів. Представлена методика може бути використана у процесі обробки інформаційних запитів під час інформаційного пошуку, даючи змогу визначити ступінь відповідності документа до інформаційних потреб користувача та підвищити пертинентність.

Використані джерела

1. *Dmytro Lande, Oleh Dmytrenko, Using Part-of-Speech Tagging for Building Networks of Terms in Legal Sphere // Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021). Volume I: Main Conference Lviv, Ukraine, April 22-23, 2021. CEUR Workshop Proceedings (ceur-ws.org). - Vol-2870. - pp 87-97. ISSN 1613-0073.*