

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК ТА
ТЕХНОЛОГІЙ



МІЖНАРОДНА
НАУКОВО-ТЕХНІЧНА КОНФЕРЕНЦІЯ
«ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ
ЛІНГВІСТИЧНОГО АНАЛІЗУ»

24-25 жовтня 2023 року

Тези доповідей

Київ 2023

Міжнародна науково-технічна конференція «Інтелектуальні технології лінгвістичного аналізу»: Тези доповідей. – К.: НАУ, 2023. – 87с.

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., професор завідувач кафедри кібербезпеки Національного технічного університету України «Київський політехнічний інститут» (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 комп'ютерних наук та технологій Національного авіаційного університету (протокол №8 від 16.10.2023 р.)

© Національний авіаційний університет, 2023

*Дмитро Ланде¹, д.т.н.,
Анатолій Фегер¹,
Dimitri Busch², PhD,*

¹КПІ ім. Ігоря Сікорського

²Fraunhofer Information Center for Planning and Building

СЕМАНТИЧНЕ ІНДЕКСУВАННЯ МАСИВІВ РЕФЕРАТИВНОЇ ІНФОРМАЦІЇ НА ОСНОВІ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ

Сьогодні проходить чергова технологічна революція — штучний інтелект стає загальнодоступним, завдяки таким системам, як GPT (від англ. — Generative Pre-trained Transformer), що можуть генерувати текстовий контент, який наближається до людського рівня [1]. Метою роботи є представлення можливостей обробки реферативної інформації із застосуванням систем генеративного штучного інтелекту (ГСШ), зокрема, ChatGPT, для вирішення задач формування семантичних мап, семантичного індексування, аналізу та візуалізації, що дозволяє розглядати такі системи як корисний аналітичний інструмент. Найчастіше при використанні ГСШ сутності можна витягувалися безпосередньо з навченої моделі. Але коли досліджується реферативна інформація для її семантичного індексування має сенс вставляти самі реферати в запити до ГСШ (промпти) [2]. Як приклад, для індексування окремих рефератів, авторами були застосовані промпти до системи ГСШ, що містили текст рефератів, такого вигляду:

Output 20 pairs of the most related entities from the text. Each entity should be described in no more than 3 words. Format of the output: 'entity 1; entity 2'. Each pair is displayed on a new line. Text: Building Information Modeling has been used to analyze as well as increase the energy efficiency of the buildings ...

У результаті виконання наведеного промпту ГСШ видає пари сутностей, які стають основою для семантичної індексації:

Machine Learning
eBIM-GNN; energy analysis
building information modeling; energy efficiency
buildings; deconstruction
retrofitting; energy savings...

