

РЕЦЕНЗІЯ

на дисертаційну роботу

Гайко Світлани Іванівни

на тему «**Онтолого-керовані засоби представлення неструктурованої інформації**»,

представлену на здобуття ступеня доктора філософії

в галузі знань Інформаційні технології

за спеціальністю **122 Комп'ютерні науки**

Актуальність теми дисертації.

Дисертаційна робота Гайко Світлани Іванівни «Онтолого-керовані засоби представлення неструктурованої інформації», присвячена вирішенню актуальної науково-технічної задачі систематизації природномовних видань в глобальному середовищі шляхом розробки моделі, методу і засобів їх трансдисциплінарного представлення та консолідації.

Актуальність теми доведена тим, що для обробки природних мов, необхідно вирішувати складні системні задачі, пов'язані із трансдисциплінарною консолідацією різностильових документів, обробленням наративів, урахуванням різних тематичних профілів та міжконтекстної зв'язності. Цей підхід дозволяє досягти якісно нового рівня при вирішенні задач автоматичного аналізу корпусів текстів, з подальшим використанням результатів в інформаційних системах підтримки управлінських рішень в різних сферах.

Робота є підсумком наукових досліджень, що виконувалися у 2013-2019 роках, з метою розробки методів, технологій та засобів автоматизованої обробки неструктурованої інформації.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

– Виконано розвиток теорії трансдисциплінарності інформації. Розроблено ряд моделей її онтологічного представлення, оцінювання, та трансдисциплінарної консолідації. Визначено рекурсивні та редуکتивні властивості різностильових документів;

– Вперше визначено модель та метод трансдисциплінарної консолідації різностильових документів;

– Вперше розроблено онтологічний класифікатор стилів, використання якого забезпечує підвищення ефективності семантичного оброблення наративів та документів різних тематичних профілів, а також реалізує виявлення їх міжконтекстної зв'язності на основі формування взаємозв'язаних онтологій, які відображають їх структуру;

– Вперше створено модель онтолого-керованої ідентифікації інформації, яка за рахунок онтології стилів забезпечила підвищення ефективності процесу створення бази правил редукції;

– Розроблена ІТ-технологія формування онтологій видань на основі репрезентації їх таксономій, зав'язків між їх контекстами та множинного представлення документів, що забезпечує ефективний пошук інформації у природномовних джерелах;

– Розроблено шкалу формалізованості різностильових документів на основі рекурсивної редукції;

– На основі онтології задачі вибору розроблено спосіб оцінки сучасних засобів представлення електронних образів видань, що дозволило виявити переваги Рекурсивного Редуктора для автоматизованої обробки природномовних текстів.

Достовірність наукових результатів забезпечується застосуванням методів системологічної класифікації, теорії складних систем, системного аналізу, теорії множин, теорії графів, лямбда-числень, алгебро-логічного та аксіоматичного методів, шаблони проектування та об'єктно-орієнтований аналіз.

Наукові дослідження були виконані здобувачем в Інституті телекомунікацій і глобального інформаційного простору у відділі інформаційної безпеки під керівництвом доктора технічних наук, професора Стрижака Олександра Євгенійовича.

Отже, в дисертаційній роботі поставлене наукове завдання, забезпечити онтолого-керованого представлення неструктурованої інформації на засадах автоматичного пошуку інформації в природномовних виданнях, їх відображення у форматі онтологій та консолідації їх контекстів, виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.

За своїм змістом дисертаційна робота здобувача Гайко Світлани Іванівни повністю відповідає Стандарту вищої освіти зі спеціальності 122 Комп'ютерні

науки та напрямкам досліджень відповідно до освітньої програми Інформаційні технології.

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям технічних наук.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Гайко Світлани Іванівни є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

Мова та стиль викладення результатів.

Дисертаційна робота написана українською мовою.

Робота виконана якісно, із застосуванням логічних переходів та узагальнень, викладена науковим стилем грамотною українською мовою. Загальноприйняті термінологія та спеціальні терміни використані відповідно своєму значенню з необхідними поясненнями.

Дисертація складається з вступу, 4 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 187 сторінок, з яких 11 сторінок додатків.

У **вступі** виконано обґрунтування вибору теми дослідження, розглянуто зв'язок з науковими програмами, сформульовано мета та завдання дослідження, розкрито наукову новизну та практичну цінність.

У **першому** розділі виконано аналіз літературних джерел, що пов'язані із темою дисертаційного дослідження. Окрім цього проведено огляд існуючих засобів для обробки та представлення природномовних текстів. Розроблено критерії оцінки обраних засобів і здійснено їх ранжування. Проведений аналіз надав обґрунтування, щодо необхідності розроблення інструментів аналізу, здатних представляти результати обробки неструктурованих текстових масивів у вигляді, що задовольняв би критерію трансдисциплінарності.

Другий розділ присвячено розробленню концептуальної моделі трансдисциплінарного представлення мережових інформаційних ресурсів. Наведено обґрунтування для створення програмно-інформаційного засобу трансдисциплінарного представлення інформаційних ресурсів різних стилів. Представлено модель онтолого-керованої ідентифікації інформації, в результаті чого отримав розвиток метод рекурсивної редукції. Проведено класифікацію функціональних стилів мови. Розроблено шкалу ступеню формалізованості документу. Проведено оцінку еквівалентності документів різних стилів з точки зору автоматизованої обробки.

Третій розділ містить детальну інформацію, щодо розробленої інформаційної моделі програмної системи, яка реалізує трансдисциплінарне представлення різностильових інформаційних ресурсів, що включає опис модулів: інтерпретації онтології стилів; формування зростаючої пірамідальної мережі; класифікації вхідного тексту за стилем; інтерпретації шаблонів представлення. Розроблено модель поведінки системи, що описує: варіанти використання системи; активність системи; взаємодію системи. На основі інформаційної моделі та моделі поведінки системи трансдисциплінарного представлення різностильових інформаційних ресурсів розроблено її архітектуру, яка визначає структуру, компонентний склад і функціональність розроблених програмних засобів.

У **четвертому** розділі розроблено інструментальний засіб «Система автоматизованої обробки різностильових документів», який надає можливість витягу із множини текстових документів знань, релевантних до заданої предметної галузі, їх системно-онтологічної структуризації та формально-логічного подання. Окрім цього, проілюстровано застосування науково-технологічних та методичних засобів онтолого-керованого трансдисциплінарного представлення великих масивів слабо і неструктурованої інформації, для підтримки процесів пошуку, класифікації та вибору необхідного контенту. Створено знання-орієнтовані персональні бібліотеки та репозиторії. Виявлено критеріальні ознаки щодо оцінювання та прийняття рішень на основі методології системно-онтологічного аналізу інформаційних ресурсів. Проведено експериментальні дослідження в ході яких підтверджено ефективність використання розробленої системи.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

Оприлюднення результатів дисертаційної роботи.

Наукові результати дисертації висвітлені у 18 наукових публікаціях здобувача, серед яких: 4 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України.

Також результати дисертації були апробовані на 12 наукових фахових конференціях, з яких 1 входить до наукометричної бази SCOPUS, а також 1 авторське свідоцтво.

Всі роботи виконані на належному рівні, що доведено незалежним рецензуванням в процесі подачі матеріалів до друку, з дотримання правил академічної доброчесності та мають особистий внесок здобувача у вигляді розробленої шкали ступеню формалізованості документу, класифікації функціональних стилів мови, розробленої моделі онтолого-керованої ідентифікації інформації, інформаційної моделі програмної системи та її архітектури, наведені прикладів застосування науково-технологічних та

методичних засобів онтолого-керованого трансдисциплінарного представлення великих масивів слабко і неструктурованої інформації, для підтримки процесів пошуку, класифікації та вибору необхідного контенту.

Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

Недоліки та зауваження до дисертаційної роботи.

Дисертаційна робота має наступні недоліки та зауваження:

1. В роботі бажано більш подрібніше надати описання того, скільки та звідки експертів було залучено для експертної оцінки підстилів, про які іде мова у другому розділі, на сторінці 90 та у Додатку Г.

2. В першому розділі при аналізі засобів для обробки природних мов наведено 26 систем обробки природномовних текстів, більшість з яких маловідомі, але серед них чомусь відсутні широковідомі системи головних вендорів ІТ ринку як NLTK (Python), Amazon Comprehend, Google Cloud NLP API, SAS Text Miner, Clarabridge.

3. У списку абревіатур відсутнє скорочення, що містяться в роботі, а саме ПМ, ПШ, ІС, ПМТ та деякі інші.

3. В роботі зустрічаються деякі неточності в оформленні, наприклад, в першому розділі у висновках невірна нумерація пунктів задач дослідження, сторінка 50. Окрім цього список літератури не завжди вірно оформлено, наприклад, джерело 10, сторінка 164, не має вказання назви періодичного видання.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу.

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Гайко Світлани Іванівни на тему «Онтолого-керовані засоби представлення неструктурованої інформації» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для Інформаційних технологій. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії»,

затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Гайко Світлана Іванівна заслуговує на присудження ступеня доктора філософії в галузі знань Інформаційні технології за спеціальністю 122 Комп'ютерні науки.

Рецензент:

Провідний науковий співробітник
ІТІПІ НАН України, д.т.н., доцент



Олександр ТЕРЕНТЬЄВ



«26» вересня 2023 року

Підпис провідного наукового співробітника ІТІПІ НАН України, д.т.н.,
доцента Олександра Терентьєва засвідчую.

Вчений секретар ІТІПІ НАН України,
к.т.н., с.н.с.



Вікторія КЛИМЕНКО