ВІДГУК ОФІЦІЙНОГО ОПОНЕНТА

доктора технічних наук, професора **ЛАНДЕ Дмитра Володимировича**

на дисертаційну роботу **ЛІТВІН Анни Андріївни**

«Методи та засоби аналізу і синтезу природномовних текстів на основі онтологічної моделі предметної області в аспекті створення діалогових систем», подану на здобуття наукового ступеня доктора філософії

за спеціальністю 122 Комп'ютерні науки

1. Актуальність теми дисертаційної роботи

Актуальність теми дослідження Літвін А. А. беззаперечна, враховуючи стрімкий розвиток технології автоматизованої обробки, зокрема аналізу і синтезу, природномовних текстів, та їх широке застосування у багатьох сферах людської діяльності. Вдало вибрано напрям дослідження, який є надзвичайно важливим для підвищення ефективності систематизації і опрацювання великих обсягів текстових даних у сучасну інформаційну епоху.

2. Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційна робота виконана в рамках науково-дослідних робіт 159/01/0245 «Трансдисциплінарна інтелектуальна інформаційно-аналітична система супроводження процесів реабілітації при пандемії (TISP)», 2020 – 2021 рр., ДКР «Сфера» (Державний контракт № 181/18/3, 2019 р.), «Створення методів та технологічних засад формування інтерактивних баз знань» (№ 0117U000005, 2021 р.).

3. Основні результати та їх наукова новизна

У роботі розглянуто три основні класи задач тісно пов'язані між собою: семантичний аналіз природномовних текстів, автоматизована побудова онтологічних графових баз знань і синтез осмисленого тексту на базі

онтологічної моделі, для розв'язання яких запропоновано відповідні оригінальні методи і розроблено їх програмну реалізацію.

Наукова новизна роботи полягає в розробці оригінальних моделей і засобів семантичного аналізу природномовних текстів, зокрема мовою флективного типу, для автоматизованого створення онтологічних баз знань і формальних запитів до них, а також синтезу природномовних відповідей на базі отриманих результатів запитів. Значним науковим доробком є розроблення нових методів поєднання великих мовних моделей і онтологокерованого підходу до їх метанавчання, які підвищують їх ефективність і релевантність при роботі із спеціалізованими предметними областями і можуть бути застосовані в різних прикладних галузях.

4. Обґрунтованість та достовірність отриманих результатів

Ступінь обґрунтованості основних положень і висновків дисертації високий. Формулювання та детальний розгляд вирішуваних задач досягнуто завдяки аналізу значної кількості наукової сучасних наукових джерел (із них більшість англомовні). Матеріали дисертації доповідалися на 8 міжнародних конференціях та опубліковані у рецензованих академічних виданнях, в тому числі таких, що входять до Scopus та Web of Science. У роботі отримано числові значення формальних критеріїв оцінки якості функціонування розроблених програмних систем.

Основні наукові положення та висновки сформульовані відповідно до поставлених завдань дисертаційної роботи.

5. Практичне значення отриманих результатів

Робота має практичне значення у розробці методів побудови онтології на основі природномовних текстів, включаючи структуровані й неструктуровані дані. Створені онтології відповідають стандартам і використовуються у якості баз знань діалогових і довідкових систем в різних

галузях, що також використовують розроблені у роботі методи семантичного аналізу і синтезу природномовних текстів. Розроблена система OntoChatGPT залучає великі мовні моделі, наприклад API ChatGPT, і використовує онтологічну базу для мета-навчання, що дозволяє як підвищити ефективність роботи великої мовної моделі у спеціалізованих предметних областях при суттєвій економії як обчислювальних ресурсів та зменшенні необхідності роботи з ручного створення розмочених великих наборів даних.

6. Повнота викладення наукових положень та висновків в опублікованих працях

Усі положення й висновки, наявні в дисертаційній роботі, науково обґрунтовані, достовірні, логічно випливають із результатів проведеного дослідження. Повнота викладення наукових положень і висновків в опублікованих працях знаходиться на високому рівні, що дозволяє отримати повне уявлення про здійснену роботу та її результати.

7. Недоліки дисертації щодо її змісту і оформлення

По дисертації можна зробити такі зауваження.

1) У анотації до дисертації не згадуються приклади практичного втілення отриманих наукових результатів, та оцінка якості роботи розроблених програмних засобів, хоча цим питанням приділено значну увагу у основному тексті роботи.
2) Схема на рисунку 2.7 перевантажена зв'язками, що утруднює її розуміння.
3) На рисунку 2.8 стрілочки в'язків між сутністю "Wordls link" та, відповідно, "Implicit Words Link" і "Auxiliary Relationship" розгорнуті у протилежному напрямку.
4) Пункти 2.4.2.3 (з розділу 2) і 3.2.2.1 (з розділу 3) багато в чому дублюють один одного, включаючи формули 2.2 – 2.7 та 3.1 – 3.6.

5) У підрозділі 4.3 описується приклад архітектури мультиагентної системи. Не зовсім ясно де потім не практиці втілюється саме така архітектура, оскільки далі, у підрозділі 4.4, де описуються конкретні приклади реалізації, надається ряд дещо інших структурних схем програмних систем.
6) У розділі 5.3, де проводиться оцінка ефективності зворотного синтезу природномовних речень на основі онтологічного представлення, згадуються різні моделі векторизації тексту (uk_core_news_lg, xx_ent_wiki_sm та tf-idf), що застосовуються при оцінці косинусної подібності. Результати оцінки з використанням різних моделей порівнюються між собою. Проте самі моделі векторизації при цьому розглядаються як чорний ящик і не розкривається їх детальна внутрішня сутність і відмінності між ними.
7) У висновках зазначається, що при синтезі природномовних фраз на основі онтологічного представлення із використанням великої мовної моделі спостерігається викривлення форми і стилю фрази у порівнянні з оригіналом. На основі цього зазначається не втрата своєї актуальності системами генерації тексту на основі онтологій за правилами та шаблонами. Проте, можливо більш тонке налаштування (в тому числі і загаданих у роботі інструкцій-підказок) і подальший розвиток саме великих мовних моделей переведе до вирішення виявленої проблеми?

Підсумовуючи та аналізуючи вищенаведені зауваження, можна відмітити, що вони не стосуються основних положень дисертації, що виносяться на захист, і тому не впливають на загальну безсумнівно позитивну оцінку дисертації.

8. Рекомендації щодо впровадження результатів дисертаційного дослідження в практику

Рекомендації до впровадження результатів дисертації можуть включати розробку додаткових програмних інтерфейсів для інтеграції розроблених систем із сервісами обміну повідомленнями і спеціалізованими довідковими і інформаційно-аналітичними системами, а також адаптацію моделей під специфічні потреби різних галузей, де застосовується автоматизована обробка (семантичний аналіз і синтез) природномовного тексту.

Значна частина результатів дисертації може бути використана у навчальному процесі вищих навчальних закладів України для студентів ІТ-спеціальностей, де викладаються такі дисципліни як онтологічна інженерія, бази даних і бази знань, людино-машинна взаємодія, теорія прийняття рішень, тощо.

9. Висновок на відповідність дисертації встановленим вимогам

Дисертаційна робота Літвін А.А. на тему: «Методи та засоби аналізу і синтезу природномовних текстів на основі онтологічної моделі предметної області в аспекті створення діалогових систем» на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» є завершеним вирішенням наукової проблеми – автоматизованого семантичного аналізу і синтезу природномовного тексту мовою флективного типу, яка супроводжувалась як комплексним аналізом існуючих методів, так і створенням нових. Ефективність запропонованих методів та алгоритмів перевірялась на основі аналізу обчислювальних експериментів. Отримані наукові результати сприяють розвитку методології розв'язання задач обробки природномовного тексту при автоматизованій побудові графових онтологічних баз даних і баз знань, формальних запитів до них і формуванню відповідей на основі результатів їх виконання.

Зважаючи на вищесказане, дисертація Літвін А. А. за актуальністю, обсягом і рівнем проведених досліджень, науковою новизною і практичною значимістю відповідає вимогам «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 №44 від 12 січня 2022 року, а її авторка заслуговує присудження ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки».

Офіційний опонент:

Доктор технічних наук, професор,
завідувач кафедри інформаційної
безпеки Навчально-наукового
фізико-технічного інституту
КПІ ім. Ігоря Сікорського _____ Дмитро ЛАНДЕ

(підпис)

Підпис гр. _____
ЗАСВІДЧУЮ
Відділ кадрів та діловодства
_____ (_____)
підпис пр-ще