

## РЕЦЕНЗІЯ

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 обраної теми дисертації.

Використання діалогових та інтерактивних довідкових систем останнім часом розширюється в багатьох сферах, а процеси семантичного аналізу і синтезу природномовних текстів відіграють ключову роль, як при формуванні відповідних баз знань для таких систем, так і безпосередньо для їх ефективної роботи. Тому в сучасних умовах тематику дисертаційної роботи, присвяченої розробці методів і програмних засобів семантичного аналізу і генерації природномовного тексту (особливо застосування онтологічних підходів для великих мовних моделей), вважаю досить актуальною.

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

Дисертаційна робота виконана у відповідності з планами науково-дослідних робіт Інституту кібернетики імені В.М. Глушкова НАН України:

– проєкт № 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.);

– проєкт «Створення методів та технологічних засад формування інтерактивних баз знань» (№ 0117U000005, 2021 р.).

### 3. Основні результати та їх наукова новизна.

У роботі запропоновано нові підходи до семантичного аналізу текстів, особливо мовами флективного типу, а також до синтезу текстів на основі онтологічної моделі. Розроблено оригінальні програмні рішення для застосування цих методів для задач автоматизованого формування графових баз даних і баз знань, запитів до них мовами SPARQL і Cypher, а також формування природномовних відповідей на базі результатів виконання запитів і аналізу вхідного природномовного запиту. Також заслуговує уваги поєднання онтології та великих мовних моделей, як в аспекті метанавчання таких моделей з використанням контекстних баз знань і метаонтології, так і для задач генерації тексту на основі семантичного представлення.

Розроблена система OntoChatGPT використовує великі мовні моделі, такі як API ChatGPT, і взаємодіє з онтологічною базою знань для метанавчання та поліпшення результатів діалогових систем, що базуються на цих моделях, у спеціалізованих предметних областях. Система також використовує запропоновану у роботі техніку метанавчання з використанням онтологічної бази контекстів і інструкції підказок, для роботи з отриманими з

неї результатами. Це дозволяє значно зменшити необхідні для спеціального донавчання такої системи обчислювальні ресурси, та ручну розмітку великого набору даних, оскільки формування контекстної онтології може відбуватися автоматично.

#### **4. Практичне значення отриманих результатів**

Практична важливість даної роботи проявляється у розробці та програмній реалізації методів побудови онтології на основі як структурованих, так і неструктурованих природномовних текстів. Цей підхід дозволяє створювати бази знань в різноманітних предметних галузях. Онтології, розроблені в рамках цієї роботи, відповідають стандартам та можуть ефективно використовуватися з графовими системами управління базами даних, такими як Apache Jena Fuseki та Neo4J, для розробки природномовних діалогових систем.

Розроблені на базі запропонованих моделей аналізу та синтезу природномовні діалогові і довідкові системи із онтологічними базами знань мають важливе значення у різних галузях, включаючи, зокрема, реабілітаційну медицину.

#### **5. Ступінь обґрунтованості основних положень та висновків дисертації.**

Основні наукові положення, висновки, рекомендації дисертаційної роботи мають належне методологічне обґрунтування. Проведено необхідні обчислювальні експерименти для оцінки ефективності методів та роботи програмних засобів, що їх використовують.

Основні результати дисертації детально висвітлені у публікаціях автора, в тому числі фахових виданнях та працях конференцій. Постановка задачі, методи та основні висновки детально подані в роботі.

#### **6. Повнота викладення наукових положень та висновків в опублікованих працях.**

Основні результати роботи викладено в 12 наукових публікаціях, що включають також англomовні видання, що індексуються у міжнародній наукометричній базі Scopus. Результати також апробовано на п'яти міжнародних наукових конференціях та відображені у відповідних друкованих матеріалах цих конференцій.

Наукові положення і висновки в опублікованих працях дозволяють науковій спільноті отримати повне уявлення про результати наукової роботи.

#### **7. Недоліки дисертації щодо її змісту і оформлення.**

Суттєвих недоліків щодо змісту і оформлення дисертації не вбачаю. Наведу лише декілька зауважень.

1) У роботі має місце повторення викладеного матеріалу у різних розділах, зокрема, наприклад формули 2.2 – 2.7 і 3.1 – 3.6 є ідентичними. Також має місце повторення опису онтологій знань між розділом 2 (де йдеться про їх автоматизоване формування) і розділом 4 (де йдеться про їх використання у якості баз знань діалогових систем).

2) Розроблена концепція виведення похідних семантичних категорій із аксіоматичного набору викладена надто стисло, хоча подається як один із важливих результатів. Не зовсім зрозуміло за якими критеріями виводяться похідні категорії і чи всі вони можуть бути ототоженені із поняттями навколишнього світу.

3) Підрозділ 1.3 оглядового розділу написано доволі широко і дещо близько скоріш до науково-популярного стилю.

4) В літературному огляді автор фокусується здебільшого на сучасних підходах до роботи з природно мовними текстами, що є загалом вірно, проте слід було б також додати коротку історичну довідку про розвиток відповідних методів, алгоритмів і програмних систем, що їх реалізують.

Наведені зауваження в жодному разі не впливають на загальну позитивну оцінку дисертаційної роботи.

#### **8. Рекомендації щодо впровадження результатів дисертаційного дослідження в практику.**

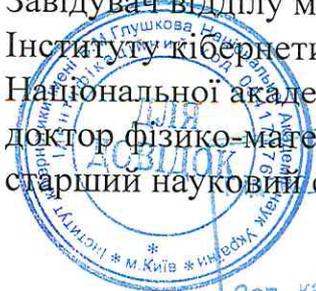
Розроблені у роботі методи та моделі можуть бути використані на практиці для створення довідкових і інформаційно-аналітичних систем у широкому колі предметних областей. Важливо, що створене програмне забезпечення вже розгорнуто на сервері і має здатність до інтеграції з іншими програмними системами у вигляді набору сервісів.

#### **9. Відповідність дисертації встановленим вимогам.**

Дисертаційна робота Літвін Анни Андріївни «Методи та засоби аналізу і синтезу природномовних текстів на основі онтологічної моделі предметної області в аспекті створення діалогових систем» є завершеною науково-дослідною роботою, виконаною на належному методичному та науковому рівні.

Зважаючи на вищесказане, дисертація Літвін А.А. за актуальністю, науковою новизною, обґрунтованістю наукових положень, теоретичним і практичним значенням отриманих результатів повністю 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».

Завідувач відділу методів системного моделювання  
Інституту кібернетики імені В. М. Глушкова  
Національної академії наук України,  
доктор фізико-математичних наук, *В. Пепеляєв*  
старший науковий співробітник



*Володимир Пепеляєв*  
Зав. канц.  
ІК НАН України

Володимир ПЕПЕЛЯЄВ