

# Облікова картка дисертації

## I. Загальні відомості

Державний обліковий номер: 0824U000807

Особливі позначки: відкрита

Дата реєстрації: 05-02-2024

Статус: Наказ про видачу диплома

Реквізити наказу МОН / наказу закладу:



## II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Літвін Анна Андріївна

2. Anna A. Litvin

Кваліфікація: 122

Ідентифікатор ORCID ID: 0000-0002-5648-9074

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 122

Назва наукової спеціальності: Комп'ютерні науки

Галузь / галузі знань: інформаційні технології

Освітньо-наукова програма зі спеціальності: Комп'ютерні науки

Дата захисту: 16-02-2024

Спеціальність за освітою: Інженерія програмного забезпечення

Місце роботи здобувача: Інститут кібернетики імені В. М. Глушкова Національної академії наук України

Код за ЄДРПОУ: 05417176

Місцезнаходження: проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

### **III. Відомості про організацію, де відбувся захист**

**Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради):** ДФ 26.194.011

**Повне найменування юридичної особи:** Інститут кібернетики імені В. М. Глушкова Національної академії наук України

**Код за ЄДРПОУ:** 05417176

**Місцезнаходження:** проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

**Форма власності:** Державна

**Сфера управління:** Національна академія наук України

**Ідентифікатор ROR:**

### **IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію**

**Повне найменування юридичної особи:** Інститут кібернетики імені В. М. Глушкова Національної академії наук України

**Код за ЄДРПОУ:** 05417176

**Місцезнаходження:** проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

**Форма власності:** Державна

**Сфера управління:** Національна академія наук України

**Ідентифікатор ROR:**

### **V. Відомості про дисертацію**

**Мова дисертації:** Українська

**Коди тематичних рубрик:** 20.23.25, 28.23.13, 28.23.37, 20.54.06

**Тема дисертації:**

1. Методи та засоби аналізу і синтезу природномовних текстів на основі онтологічної моделі предметної області в аспекті створення діалогових систем
2. Methods and means of natural language texts analysis and synthesis based on the ontological model of the subject area in the aspect of dialogue systems development

**Реферат:**

1. Дисертація присвячена дослідженню та розробці методів та програмних засобів синтактико-семантичного аналізу і синтезу природномовних текстів на основі онтологічної моделі предметної області, орієнтованих насамперед на мови флективного типу, зокрема, українську. Серед завдань роботи були: 1) розробка методів аналізу природно-мовного тексту прийнятних для мов флективного типу, зокрема, української, у двох основних аспектах призначення: автоматичне створення графових баз знань онтологічного типу на основі природно-мовного тексту або набору таких текстів; аналіз коротких фраз і запитів користувачів для визначення намірів і іменованих сутностей, що їх конкретизують, висловлених у них, необхідних для побудови формальних запитів до бази знань; 2) Створення гнучких шаблонів формальних запитів на мовах

SPARQL та Cypher і програмних засобів для роботи з ними, необхідних для отримання релевантної інформації з бази знань; 3) розробка методів генерації осмислених природно-мовних відповідей і їх представлення користувачеві із залучанням як підходів, заснованих на правилах, так і з залученням великих мовних моделей; 4) практичне створення природно-мовних діалогових і довідкових систем, які втілюють розроблені підходи до аналізу і синтезу тексту із залученням баз знань онтологічного типу; 5) проведення аналізу роботи і тестування розроблених діалогових систем, визначення формальних критеріїв якості їх функціонування. Основні наукові результати роботи полягають у наступному: 1) Отримала подальший розвиток теорія системи семантичних відношень основною відмінністю якої є виведення значного розмаїття семантичних категорій з аксіоматичного набору понять верхнього рівня. 2) Вперше створено метод отримання формальних запитів мовами SPARQL і Cypher на основі результатів аналізу природномовних текстів, адаптованого до мов флективного типу, зокрема, української. Створення запитів відбувається з використанням гнучких шаблонів. Особливість підходу полягає в тому, що формальний запит автоматично будується із блоків шаблонів (основних і допоміжних), що є такими, що налаштовуються, відповідно визначених семантичних категорій визначених у аналізованому тексті, та сутностей, що їх конкретизують; 3) Отримали подальший розвиток методи автоматичної побудови онтологій на основі природномовного тексту, в тому числі не розміненого, на базі глибокого синтактико-семантичного аналізу. Запропоновані способи є ефективними як для мов флективного типу, зокрема, української, так і для англійської мови; 4) Отримали подальший розвиток методи для генерації відповідей за допомогою шаблонів гнучкої структури на основі набору понять, отриманих шляхом виконання формального запиту до бази знань і слів, відібраних із вхідної фрази користувача. Представлений метод адаптовано для мов флективного типу, зокрема, української. Шаблон визначає місця підстановки окремих слів та їх груп з вихідного речення та їх словоформи, додаткові текстові вставки з можливістю їх варіювання, місця підстановки результатів запиту та обмеження, що накладаються на них. Раніше таких підходів до формування відповідей мовами флективного типу не існувало. Шаблони для формування відповідей мають гнучку структуру і автоматично адаптуються згідно результатів як аналізу вхідного тексту, так і результатів виконання запиту. Великі мовні моделі добре справляються з задачею генерації тексту, але вони мало пристосовані саме до обробки і представлення результатів формальних запитів до бази знань онтологічного типу. Це підтверджено наведеними у роботі експериментальними даними зі зворотного синтезу природномовних речень на основі представлених у OWL-онтології (отриманої із тексту) понять і семантичних відношень між ними; 5) Вперше розроблений підхід, який базується на застосуванні структурованих інструкцій-підказок, що утворюються з використанням мета-онтології. Інструкції передаються великій мовній моделі, для виконання задач аналізу вхідного природномовного тексту (визначення висловлених намірів і іменованих сутностей). Далі з використанням інструкції, сформованої відповідно визначеним намірам і іменованим сутностям, великій мовній моделі передаються контексти, відібрані із бази знань предметної області за допомогою відповідних формальних запитів для отримання релевантних відповідей. Результати дисертації є суттєвим внеском у методологію процесів семантичного аналізу і синтезу природномовного тексту в аспекті автоматизованої побудови графових баз знань онтологічного типу та автоматизованого створення запитів до них.

2. The dissertation is dedicated to the research and development of methods and software tools for syntactic-semantic analysis and synthesis of natural language texts based on the ontological model of the subject area, focused primarily on languages of the inflectional type, in particular, Ukrainian. Among the tasks of the work were: 1) development of natural language text analysis methods acceptable for languages of the inflectional type, in particular, Ukrainian, in two main aspects of the purpose: automatic creation of graph knowledge bases of the ontological type based on natural language text or a set of such texts; analysis of short phrases and requests of users to determine expressed intentions and named entities that specify them, which is necessary for formal queries to the knowledge base automatic creation; 2) creation of flexible templates of formal SPARQL and Cypher queries and the sufficient software tools, necessary for obtaining relevant information from the knowledge base; 3) development of methods for meaningful natural-language responses generating and presenting them to the user involving both rule-based approaches and large language models; 4) practical creation of natural-language

dialogue and reference systems that embody developed approaches to text analysis and synthesis with the involvement of ontological-type knowledge bases; 5) analysis of work and testing of developed dialog systems to determine formal quality criteria for their functioning. The main scientific results of the work are as follows: 1) The theory of the semantic relations system has received further development. Its main difference is the derivation of a significant variety of semantic categories from the axiomatic set of upper-level concepts. 2) A method for formal queries obtaining in the SPARQL and Cypher languages was firstly created, which is based on the results of the analysis of natural language texts, adapted to languages of the inflectional type, in particular, Ukrainian. Requests are created using flexible templates. The peculiarity of the approach is that the formal query is automatically built from blocks of templates (main and auxiliary), which are customizable according to the defined semantic categories defined in the analyzed text, and the entities that specify them; 3) Methods were further developed for automatic ontologies construction based on natural language text, including unmarked text, on the basis of deep syntactic-semantic analysis. The proposed methods are effective both for languages of the inflectional type, in particular, Ukrainian, and for English; 4) Methods were further developed to generate responses using flexible structure templates based on a set of concepts obtained by executing a formal knowledge base query and words selected from the user's input phrase. The presented method is adapted for languages of the inflectional type, in particular, Ukrainian. The template defines the placeholders for individual words and their groups obtained from the original sentence and their word forms, additional text inserts with the possibility of their variation, substitution places of the query results and restrictions imposed on them. Previously, such approaches to the formation of answers in languages of the inflectional type did not exist. The templates have a flexible structure and automatically adapt according to the results of both input text analysis and the results of query execution. Large language models cope well with the task of text generation, but they are poorly adapted to the processing and presentation of the results of formal queries to the ontology-type knowledge base, which was confirmed by the experiments presented in the dissertation that included the reverse synthesis of natural language sentences based on the concepts presented in the OWL ontology (obtained from the text) and the semantic relations between them; 5) For the first time, an approach was developed, which is based on the application of structured prompts formed using meta-ontology. The prompts are transmitted to a large language model to make it perform the tasks of analyzing the incoming natural language text (determining expressed intentions and named entities). Next, using an instruction formed according to defined intentions and named entities, contexts selected from the knowledge base of the subject area are passed to the large language model using appropriate formal queries to obtain relevant answers. The results of the dissertation are a significant contribution to the methodology of the processes of natural language text semantic analysis and synthesis in the aspect of automated construction of graph knowledge bases of the ontological type and automated creation of queries to them.

**Державний реєстраційний номер ДіР:** 0117U000005,

**Пріоритетний напрям розвитку науки і техніки:** Інформаційні та комунікаційні технології

**Стратегічний пріоритетний напрям інноваційної діяльності:** Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

**Підсумки дослідження:** Нове вирішення актуального наукового завдання

**Публікації:**

- Palagin O., Kaverinskiy V., Litvin A., Malakhov K. OntoChatGPT Information System: Ontology-Driven Structured Prompts for ChatGPT Meta-Learning. International Journal of Computing. 2023. 22(2). P. 170-183.
- Litvin A., Palagin O., Kaverinsky V., Malakhov K. Ontology-driven development of dialogue systems. South African Computer Journal. 2023. 35(1). P. 37-62.
- Litvin A.A., Velychko V.Y., Kaverynskiy V.V. A dialogue system based on ontology automatically built through a natural language text analysis. Problems in programming. 2022. № 3-4. С. 196-207

- Litvin A.A., Velychko V.Y., Kaverynskyi V.V. Tree-based semantic analysis method for natural language phrase to formal query conversion. Radio Electronics, Computer Science, Control. 2021. Vol. 2. P. 105–113.
- Литвин А.А., Величко В.Ю., Каверинский В.В. Метод получения информации из онтологии на основе анализа фразы на естественном языке. Проблемы програмування. 2020. № 2–3. С. 322–330.

**Наукова (науково-технічна) продукція:** технології; методи, теорії, гіпотези

**Соціально-економічна спрямованість:** підвищення продуктивності праці; поліпшення якості життя та здоров'я населення, ефективності діагностики та лікування хворих; підвищення автоматизації виробничих процесів; забезпечення промисловості чи населення новим видом інформаційно-комунікаційних послуг

**Охоронні документи на ОПВ:**

Комп'ютерні програми

Свідоцтво про реєстрацію авторського права на твір № 96125

**Впровадження результатів дисертації:** Планується до впровадження

**Зв'язок з науковими темами:**

## **VI. Відомості про наукового керівника/керівників (консультанта)**

**Власне Прізвище Ім'я По-батькові:**

1. Палагін Олександр Васильович

2. Oleksandr V. Palagin

**Кваліфікація:** д. т. н., професор, 05.13.05

**Ідентифікатор ORCID ID:** 0000-0003-3223-1391

**Додаткова інформація:** <https://www.scopus.com/authid/detail.uri?authorId=6701857667>

**Повне найменування юридичної особи:** Інститут кібернетики імені В. М. Глушкова Національної академії наук України

**Код за ЄДРПОУ:** 05417176

**Місцезнаходження:** проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

**Форма власності:** Державна

**Сфера управління:** Національна академія наук України

**Ідентифікатор ROR:**

## **VII. Відомості про офіційних опонентів та рецензентів**

**Офіційні опоненти**

**Власне Прізвище Ім'я По-батькові:**

1. Ланде Дмитро Володимирович

2. Dmytro V. Lande

**Кваліфікація:** д. т. н., професор, 05.13.06

**Ідентифікатор ORCID ID:** 0000-0003-3945-1178

**Додаткова інформація:** <https://www.scopus.com/authid/detail.uri?authorId=19934022600>

**Повне найменування юридичної особи:** Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

**Код за ЄДРПОУ:** 02070921

**Місцезнаходження:** проспект Берестейський, буд. 37, Київ, 03056, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

**Власне Прізвище Ім'я По-батькові:**

1. Погорілий Сергій Дем'янович

2. Sergii D. Pogorilyu

**Кваліфікація:** д.т.н., професор, 05.13.05

**Ідентифікатор ORCID ID:** 0000-0002-6497-5056

**Додаткова інформація:** <https://www.scopus.com/authid/detail.uri?authorId=26428941500>

**Повне найменування юридичної особи:** Київський національний університет імені Тараса Шевченка

**Код за ЄДРПОУ:** 02070944

**Місцезнаходження:** вул. Володимирська, буд. 60, Київ, 01033, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

**Рецензенти**

**Власне Прізвище Ім'я По-батькові:**

1. Горбачук Василь Михайлович

2. Vasyl M. Gorbachuk

**Кваліфікація:** д.ф.-м.н., с.н.с., 01.05.01

**Ідентифікатор ORCID ID:** 0000-0001-5619-6979

**Додаткова інформація:** <https://www.scopus.com/authid/detail.uri?authorId=55646075900>

**Повне найменування юридичної особи:** Інститут кібернетики імені В. М. Глушкова Національної академії наук України

**Код за ЄДРПОУ:** 05417176

**Місцезнаходження:** проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

**Форма власності:** Державна

**Сфера управління:** Національна академія наук України

**Ідентифікатор ROR:**

**Власне Прізвище Ім'я По-батькові:**

1. Пепеляев Володимир Анатолійович
2. Volodymyr A. Pepelyaev

**Кваліфікація:** д. ф.-м. н., старший науковий співробітник, 01.05.02

**Ідентифікатор ORCID ID:** 0009-0009-3169-1776

**Додаткова інформація:** <https://www.scopus.com/authid/detail.uri?authorId=8967548400>

**Повне найменування юридичної особи:** Інститут кібернетики імені В. М. Глушкова Національної академії наук України

**Код за ЄДРПОУ:** 05417176

**Місцезнаходження:** проспект Академіка Глушкова, буд. 40, Київ, 03187, Україна

**Форма власності:** Державна

**Сфера управління:** Національна академія наук України

**Ідентифікатор ROR:**

## VIII. Заключні відомості

**Власне Прізвище Ім'я По-батькові  
голови ради**

Будник Микола Миколайович

**Власне Прізвище Ім'я По-батькові  
головуючого на засіданні**

Будник Микола Миколайович

**Відповідальний за підготовку  
облікових документів**

Стовба Віктор Олександрович

**Реєстратор**

УкрІНТЕІ

**Керівник відділу УкрІНТЕІ, що є  
відповідальним за реєстрацію наукової  
діяльності**



Юрченко Тетяна Анатоліївна