

В. В. Каверинський, А. А. Літвін, О. В. Палагін

ЗВОРОТНИЙ СИНТЕЗ ПРИРОДНОМОВНИХ ВИСЛОВЛЮВАНЬ НА ОСНОВІ ЇХ ОНТОЛОГІЧНОГО ПРЕДСТАВЛЕННЯ З ВИКОРИСТАННЯМ ВЕЛИКОЇ МОВНОЇ МОДЕЛІ

Стаття презентує нове рішення, що полягає у використанні спеціально розробленої структурованої інструкції-підказки для великої мовної моделі (Chat GPT). Було проведено серію експериментів із синтезу природномовних висловлювань на основі онтологічних представлень, автоматично побудованих на основі речень науково-технічного тексту за допомогою раніше розроблених програмних засобів. Ці представлення містять сутності, знайдені у тексті, та типізовані семантичні зв'язки між ними, які можуть реалізовуватися у фразах аналізованого тексту. Система зв'язків, конкретизованих набором понять, поєднується із сутністю зв'язаної частини речення, яка, своєю чергою, може бути простим реченням або частиною складного речення. Структурована інструкція-підказка для великої мовної моделі містить роз'яснення семантичних відношень між поняттями у контексті синтезу речень із онтологічного представлення, а також набір пар понять, поєднаних семантичними зв'язками, що слугують матеріалом для створення речення. Синтезовані у такий спосіб природномовні речення порівнювалися з оригіналами за показником косинусної близькості за умов різних методів векторизації, зокрема `xx_ent_wiki_sm`, `uk_core_news_lg` та `tf-idf`. Отримані показники близькості сягали 0,8193 – 0,9722 за моделлю `xx_ent_wiki_sm`, хоча у деяких випадках спостерігалось стилістичне викривлення отриманих речень. Дослідження має практичне значення для розробки діалогових інформаційних систем, що поєднують онтологічний підхід із використанням великих мовних моделей.

Ключові слова: велика мовна модель, онтологія, синтез природномовного тексту, аналіз природномовного тексту, косинусна близькість, векторизація тексту.

V. V. Kavarynskyi, A. A. Litvin, O. V. Palagin

REVERSE SYNTHESIS OF NATURAL LANGUAGE PHRASES GROUNDING ON THEIR ONTOLOGICAL REPRESENTATION USING A LARGE LANGUAGE MODEL

The presented article introduces a novel solution that uses a specially developed structured prompt for a large language model (Chat GPT). A series of experiments were carried out on synthesizing natural language phrases based on their ontological representations. These ontological representations were automatically constructed from sentences of scientific and technical texts using previously developed software tools. Such representations contain entities found in the text and typed semantic relationships between them, which can be realised in the phrases of the analysed text. The system of relationships, specified by a set of concepts, is linked with the entity of the related part of the sentence, which in turn can be a simple sentence or part of a complex sentence. The structured prompt for the large language model includes explanations of the semantic relationships between concepts in the context of sentence synthesis from ontological representation, as well as a set of pairs of concepts connected by semantic relationships, which serve as material for sentence creation. The synthesised natural language sentences were compared with the originals using the cosine similarity measure across different vectorisation methods. The obtained similarity scores ranged from 0.8193 to 0.9722 according to the `xx_ent_wiki_sm` model, although stylistic distortions of the generated sentences were observed in some cases. The research presented in this work has practical significance for the development of dialogue information systems that combine the ontological approach with the use of large language models.

Key words: large language model, ontology, natural language text synthesis, natural language text analysis, cosine similarity, text vectorization.

Вступ

Дослідження у сфері розвитку онтологій та семантичної обробки наукових даних мають велике значення для сучасної наукової спільноти, особливо в контексті швидкого зростання обсягу інформації та потреб дослідників у її ефективному використанні. Основним завданням є розроблення та впровадження технологій, що дозволять швидко й точно знаходити й обробляти наукову інформацію, а також взаємодіяти з інформаційними системами для забезпечення максимальної корисності цих даних [1, 2]. Для вирішення цих завдань були розроблені різні системи та методології обробки наукових даних, наприклад [1–3]. Вони використовують такі технології, як Semantic Web і когнітивна графіка, щоб полегшити пошук інформації та відкриття знань у цифрових бібліотеках. Дослідження у галузі цифрового здоров'я та телереабілітації також займають важливе місце у сучасній науці [3–7]. Дослідники активно працюють над розробкою та впровадженням інноваційних технологій у цій сфері, зокрема, розвитком хмарних платформ для телереабілітації пацієнтів [5, 6]. Ці платформи базуються на поєднанні штучного інтелекту та математичного моделювання, що дозволяє створювати ефективні та інноваційні підходи до реабілітації та підтримки здоров'я [5, 8].

Стан розробки проблеми. Розвиток онтологічної інженерії в поєднанні з нейронно-мережевими технологіями та штучним інтелектом відкриває нові перспективи у взаємодії з інформаційними системами та у розвитку інноваційних сервісів [8, 9]. Наприклад, у праці [9] описано інтегрований підхід, що поєднує нейронні мережі й онтолінгвістичні парадигми для підвищення ефективності інтелектуальних діалогових систем та забезпечення більш гнучкого їх адаптування до різних потреб користувачів і предметних областей [9]. Цей інтегрований підхід також підкреслює важливість метанавчання та структурованих підказок для підвищення ефективності мовної моделі.

У представленому в даній роботі дослідженні подальший розвиток отримав підхід до інтеграції великих мовних моделей та онтологічних структур знань і застосування структурованих інструкцій-підказок, зокрема в аспекті процесів відтворення природномовного тексту на основі семантичного представлення. Далі для зазначеного підходу вводиться термін зворотний синтез.

Метою дослідження є розроблення структурованої інструкції-підказки для великої мовної моделі для синтезу природномовних висловлювань на основі їхніх онтологічних представлень, а також подальше оцінювання результатів шляхом формального порівняння згенерованих фраз із оригінальними текстами за допомогою різних моделей векторизації тексту.

Спосіб побудови природномовної фрази на основі онтологічного представлення з використанням великої мовної моделі

Автоматично створена онтологія, отримана з тексту природної мови, містить сутності та семантичні зв'язки, які їх з'єднують. Ці зв'язки, задані сутностями, пов'язані з групами виразів, які, своєю чергою, пов'язані з реченнями у вхідному тексті. Отже, з таким семантичним представленням можна сформулювати речення природною мовою. Для цього завдання можна застосувати шаблонний підхід, як описано в [10]. Однак із розвитком підходів, заснованих на використанні нейронних мереж глибокого навчання, втіленням яких є великі моделі мовних трансформаторів, такі як ChatGPT, виникає необхідність дослідити їх як інструменти для синтезу речень природної мови на основі семантичних структур. Спроба в цьому напрямку була зроблена в рамках даної роботи.

За тестову онтологію використано базу знань, створену на основі тексту «Склад обчислювальної системи». Оскільки база знань управлялася графовою СУБД

Neo4J, мовою запитів для неї була обрана Cypher.

Нижче наведено текст запиту до вказаної онтології для отримання тексту конкретного речення (для порівняння результатів) за його ідентифікатором:

```
MATCH (inp:Relationship)-[:SPO]-
>(inp_type:Relationship),
  (inp:Relationship)-[:SPO]-
(linked_group:Relationship),
  (linked_group:Relationship)-[:SPO]-
>(linked_group_type:Relationship),
  (linked_group:Relationship)-[:SPO]-
(certain_words_link:Relationship),
  (certain_words_link:Relationship)-
[:SPO]->(sem_type:Relationship),
  (sem_type:Relationship)-[:SPO]-
>(w_link_type:Relationship),
  (certain_words_link:Relationship)-
[:DOMAIN]->(main_entity:Class),
  (certain_words_link:Relationship)-
[:RANGE]->(dependent_entity:Class)
WHERE
  inp_type.name = "SentenceGroups" and
  linked_group_type.name = "Groups" and
  w_link_type.name = "WordsLink" and
  ID(inp) = specify sentence ID
RETURN DISTINCT ID(inp) as id, inp.label
as text, main_entity.label as main_entity,
  dependent_entity.label as
dependent_entity, sem_type.label as sem_type;
```

Цей запит також повертає відповідний набір семантичних категорій і пов'язаних пар понять (головних і залежних) для цього речення. Результати такого запити слугували вхідними даними для завдання зворотного синтезу речення природної мови.

Відповідно до наведеного запиту із онтології повертається речення з певним ідентифікатором (specify sentence ID). Результат має містити ідентифікатор, текст, а також набір триплетів виду «основна сутність, залежна сутність і семантичний тип» щодо вказаного речення. Отримана з онтології семантична структура виявляється достатньою для побудови зв'язного природномовного речення відповідного змісту. Щоб ініціювати завдання синтезу у великій мовній моделі, такій як ChatGPT, необхідно надати відповідну інструкцію-підказку (так званий prompt). Як зазначалося в [11], для таких інструкцій бажано використовувати англійську мову. Сама інструкція структурована у форматі JSON. Відповідний текст підказки подано нижче:

```
{"Intriduction": «You are an expert in
knowledge engineering and ontologies as well
as in meaningful text generation in inflect
```

```
languages. You will be provided with data
obtained from some ontology through a query.
The ontology was made automatically basing on
the results of semantic analysis of a natural
language text. The results are pairs of
lemmatized words ("main entity" and "dependent
entity") accompanied with a name of syntactic-
semantic relationship that linked them in the
certain sentence.",
"Action to perform": «Assuming that all the
data you will be provided belong to one
sentence you are to make a try to restore the
original sentence using such a prompt.
Language of the ontology, input and output
data is Ukrainian.",
"Restrictions": «Do not put the semantic
relationships as a phrase as it given in the
sentence you generate, it will be definitely
wrong. It is just a prompt for syntactic
linking. Remember that the provided words are
lemmatized, so you are to put them in a
correct form according to other entities of
the sentence and the given syntactic-semantic
relationships of the prompt.",
"Additional data to provide": «Also provide an
estimated value of probability that the
generated sentence corresponds the intent of
the prompt given.",
"The essence of the syntactic-semantic
relationship names and meaning explanation":
{"object property": «the dependent entity
express a property or some characteristic, or
quality of the main entity. When the response
sentence generation you should use the
dependent entity as an adjective with the main
entity which is noun",
"action property": "the dependent entity
express a property or some characteristic, or
quality of the main entity which is an action.
When the response sentence generation you
should use the dependent entity as an adverb
with the main entity which is verb",
"quality change": "the dependent entity
express that the main entity may be subjected
to some quality changes, which may follow from
the other context",
"destination": «the dependent entity express
the destination of the main entity",
"object": «the object (noun) affected throw
the action expressed by the main entity",
"object / action": «the main entity performs
an action expressed by the dependent entity",
"preposition binding": «merely shows that
the main entity here in the context of the
provided sentence is to be used with the
preposition which is the dependent entity.
This means that you should use this
preposition with the main entity when the
response sentence generation",
"possession": "the dependent entity or
somewhat relates to the main entity. When
generation this usually should be expresses
using genitive case",
"equality": "the different name of the entity
or an equivalent entity",
"objective entry": "the main entity is a part
or member of the dependent entity",
"state": "a state or a constant characteristic
of the main entity if it is noun or an entity
linked to in if it is a verb",
"Input data": [{
"main entity": "some word 1",
"dependent entity": "some word 2",
"semantic relationship": «semantic category
1»,
{"main entity": "some word n",
"dependent entity": "some word n+1",
"semantic relationship": "semantic category n
}]]}
```

Дамо коротке пояснення стосовно розділів наведеної інструкції-підказки.

Розділ *"Introduction"* встановлює початкові параметри для великої мовної моделі щодо її подальшої поведінки та пропонує основні роз'яснення вхідних даних.

У розділі *"Action to Perform"* сформовано безпосереднє завдання, яке необхідно виконати.

Розділ *"Restrictions"* містить додаткові вказівки щодо сформованого вихідного тексту, які націлені на усунення неоднозначності у тлумаченні інструкції.

Розділ *"Additional Data to Provide"* слугує для вказівок моделі провести власне оцінювання якості виконання завдання.

У розділі *"The Essence of the Syntactic-Semantic Relationship Names and Meaning Explanation"* подано словник пояснень типів семантичних зв'язків і способів їх використання у побудові речення. Враховуючи значну кількість семантичних категорій за умови обмеженої кількості символів для вхідних повідомлень ChatGPT, практичний обсяг такого словника обмежений семантичними категоріями, наявними в даному реченні.

Пари сутностей і їхні відповідні семантичні зв'язки перераховані у вигляді списку словників у розділі *"Input Data"*.

На виході отримуємо сформульоване речення природною мовою та власну оцінку моделі, надану ChatGPT, щодо ймовірності того, що речення було побудовано точно та відповідає оригіналу (про вигляд його модель не знає).

Загальну схему процесу формування природномовних висловлювань на основі їхньої онтологічної репрезентації наведено на Рис. 1. Суть експерименту полягала в тому, що із тестової онтології, створеної на основі тексту «Склад обчислювальної системи», за допомогою запиту мовою Cypher видобувалися окремі речення та відповідні їм пари сутностей із семантичними категоріями, що зв'язують їх у межах даного речення. Далі, застосовуючи вищенаведену інструкцію-підказку, великій мовній моделі (ChatGPT) передавалося завдання згенерувати граматично правильне речення українською мовою на основі набору пар сутностей із заданими семантичними зв'язками. У відповідь отримано згенероване речення та зроблена моделлю оцінка ймовірності того, що речення було точно відтворено. Для тестування було використано 10 речень із зазначеного тексту.

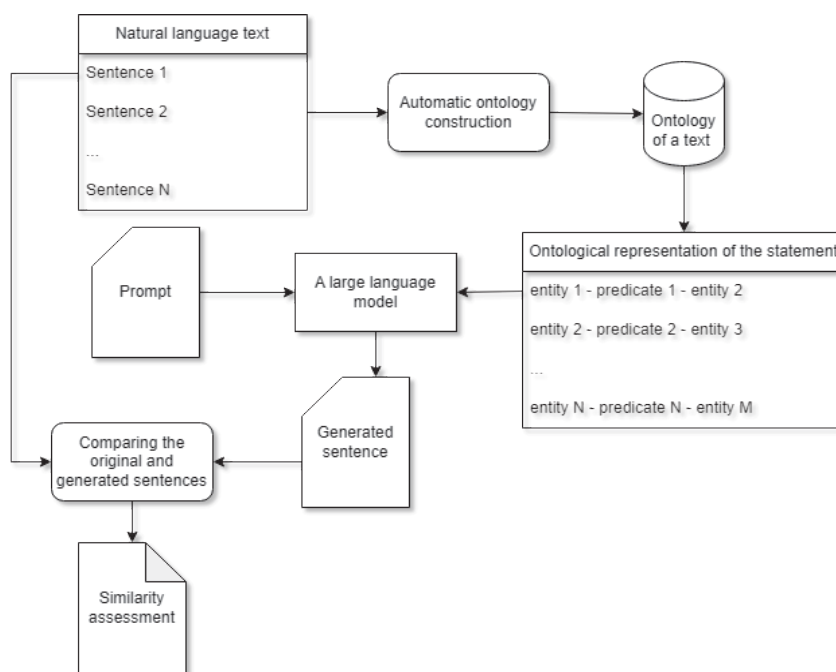


Рис. 1. Загальна схема генерації природномовних висловлювань на основі онтологічного представлення з використанням великої мовної моделі

Методика оцінювання результатів експерименту

Для порівняння подібності створеного речення до оригіналу використовувалося значення косинусної подібності. Косинусна подібність — це міра подібності між двома передгільбертовими просторовими векторами, яка використовується для вимірювання косинусу кута між ними. Отже, якщо є два вектори ознак (A і B), то косинусну подібність $\cos(\theta)$ можна представити за допомогою скалярного добутку та норми (1):

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

Косинусна подібність змінюється від 0 до 1. Це пов'язано з тим, що частота термінів (ваги tf-idf) не є від'ємною, а кут між двома векторами частоти термінів не може перевищувати 90° . Косинусна подібність є ефективним засобом оцінки, особливо для розріджених векторів, оскільки враховує лише ненульові значення. «М'який» косинус враховує подібність між парами ознак. Традиційна косинусна подібність розглядає функції у векторній моделі як незалежні або повністю окремі, тоді як «м'яка» косинусна міра визнає подібність ознак у векторній моделі. Це дозволяє узагальнити концепцію косинусної подібності та концепцію подібності об'єктів у векторному просторі.

Такі сутності, як слова, N-грами або синтаксичні N-грами, можуть демонструвати суттєву подібність, хоча формально їх можна вважати різними функціями у векторній моделі. Для N-грам або синтаксичних N-грам можна застосувати відстань Левенштейна [12]. Для обчислення «м'якої» косинусної міри вводиться матриця подібності s між функціями. Вона обчислюється за допомогою відстані Левенштейна або інших заходів подібності, таких як різні засоби подібності WordNet. Далі за цією матрицею виконується множення. Якщо є два N-вимірних вектора, a і b , «м'який» косинус обчислюється так:

$$\text{soft_cosine}_1(a, b) = \frac{\sum_{i,j}^N s_{ij} \cdot a_i \cdot b_j}{\sqrt{\sum_{i,j}^N s_{ij} \cdot a_i \cdot a_j} \cdot \sqrt{\sum_{i,j}^N s_{ij} \cdot b_i \cdot b_j}} \quad (2)$$

У формулі (2) s_{ij} представляє подібність між функціями i та j . За відсутності подібності між ознаками ($s_{ii} = 1, s_{ij} = 0$ для $i \neq j$) рівняння (2) еквівалентне загальноприйнятій формулі косинусної подібності.

Оскільки виконання математичних обчислень безпосередньо на рядках є неможливим, а обчислення такого показника, як подібність косинусів, вимагає існування векторів, тексти для обробки та аналізу проходять векторизацію. Щоб отримати векторні відображення речень, було використано бібліотеку Python spaCy і мовні моделі uk_core_news_lg (для української мови) та xx_ent_wiki_sm (багатомовна). Також застосовувався метод TF-IDF. Для обчислення значень косинусної подібності були задіяні методи, реалізовані в spaCy.

Обговорення результатів експериментів

Значення кількісних оцінок, що характеризують близькість згенерованих речень до оригіналу за різних методів векторного представлення, подані в Табл. 1. Крім того, надається «суб'єктивна» оцінка ймовірності точного відтворення від ChatGPT. Важливо зазначити, що ця оцінка не може вважатися цілком об'єктивним показником, а радше є орієнтиром та оцінкою самокритики моделі GPT. Із представлених результатів видно, що кількісна оцінка косинусної подібності значно залежить від моделі векторизації, застосованої до аналізованих текстів.

Результати показують, що мовні моделі xx_ent_wiki_sm і uk_core_news_lg забезпечують досить високі значення косинусної подібності (0,87 і 0,81 відповідно). Тим часом більш простий метод векторизації на основі tf-idf дає значно нижчі середні значення та більший діапазон варіації. Модель xx_ent_wiki_sm (багатомовна) демонструє вузький діапазон варіації та відносно високе значення середньої косинусної подібності. Зменшення середнього значення із використанням моделі uk_core_news_lg (для української мови) пояснюється більшою варіацією вниз. Однак максимальні отримані значення для цих двох моделей досить близькі – засто-

сування моделі `uk_core_news_lg` у деяких випадках призводить до значно нижчого значення косинусної подібності. Порівняння показників косинусної подібності, отриманих за допомогою моделей векторизації `xx_ent_wiki_sm` і `uk_core_news_lg`, показано на Рис. 2 (а), що свідчить про відсутність будь-якої суттєвої кореляції між ними. Тобто, ці моделі дещо по-різному сприймають текст природної мови.

Аналіз безпосередньо згенерованих речень виявив, що під час векторизації за допомогою моделі `uk_core_news_lg` знижені оцінки косинусної подібності виникають у випадках генерації складних, розгалужених речень, тоді як вихідне речення є значно простішим, хоча семантично близьким.

Таблиця 1

Кількісні оцінки якості зворотного синтезу речень

Власна оцінка від ChatGPT		Косинусна подібність					
		Модель векторизації <code>xx_ent_wiki_sm</code>		Модель векторизації <code>uk_core_news_lg</code>		Модель векторизації <code>tf-idf</code>	
Середнє значення	Інтервал значень	Середнє значення	Інтервал значень	Середнє значення	Інтервал значень	Середнє значення	Інтервал значень
0.845 ±0.037	0.75 – 0.90	0.8716 ±0.0335	0.8193 – 0.9722	0.8108 ±0.1224	0.4067 – 0.9653	0.2927 ±0.1718	0.0607 – 0.7745

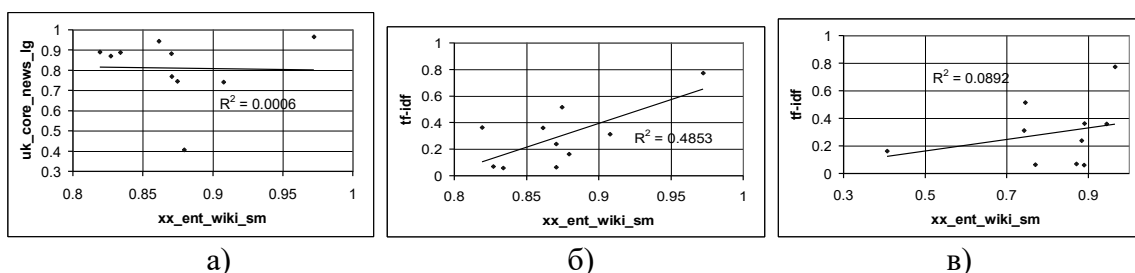


Рис. 2. Порівняння значень косинусної подібності вихідних і сформованих речень після векторизації тексту за різними моделями:

- а) `xx_ent_wiki_sm` / `uk_core_news_lg`;
- б) `xx_ent_wiki_sm` / `tf-idf`;
- в) `uk_core_news_lg` / `tf-idf`.

Модель `xx_ent_wiki_sm` видається менш чутливою до таких проявів і забезпечує вищу оцінку косинусної подібності. Однак, як видно з графіка, чіткої кореляції між моделями немає. Бувають випадки, коли `uk_core_news_lg` дає вищу оцінку косинусної подібності, а `xx_ent_wiki_sm`, відповідно, меншу. Аналіз свідчить про те, що в таких випадках, незважаючи на збереження лексично та синтаксично близької до оригіналу структури, зміст був дещо спотворений.

Існує помітна кореляція в оцінках косинусної подібності між моделлю `xx_ent_wiki_sm` і оцінками, отриманими для методу `tf-idf`. Найвищі показники за цими методами були отримані у разі практично повного збігу речень. Що ж до най-

нижчих значень оцінки, то метод `tf-idf` також виявився чутливим до деформованих речень, що знижувало значення метрик. Модель `uk_core_news_lg` слабо корелює з методом `tf-idf`. Таким чином, якщо форма представлення не має великого значення, а акцент робиться на змісті, можна застосувати векторизацію за допомогою моделі `xx_ent_wiki_sm`.

Водночас модель `uk_core_news_lg` чутлива як до змісту, так і до зміни форми, що робить її придатною для отримання жорсткішого та чутливішого порівняння. Метод `tf-idf` дуже чутливий до зміни форми, але менш вправний у розпізнаванні подібності вмісту.

Подальшим розвитком наведених досліджень може стати вдосконалення ві-

дповідних інструкцій-підказок для великих мовних моделей, у тому числі не просто для відтворення вихідного тексту за онтологічною моделлю, а й логічного висновку із поданої інформації, що наблизить нас до вирішення глобальної задачі онтологічного підходу – можливості отримання нових знань [2, 3]. Також перспективним видається можливість розгляду альтернативних GPT моделей, у тому числі автономних малих мовних моделей і, можливо, архітектури процесорів, запропонованої у роботах [13, 14]. Зрозуміло, що це потребує масштабних досліджень, пов'язаних із самою природою знань (оцінювання новизни, логікові виведення вторинних знань на основі первинних, логіко-семантичне порівняння контекстів і т.п.).

Висновки

Порівняння оцінок, отриманих різними методами, та візуальний перегляд результатів експерименту дозволяє підсумувати, що запропонований підхід до генерації природних речень українською мовою на основі їх онтологічної репрезентації за допомогою великої мовної моделі здатний передати загальний зміст і сенс оригінальної фрази, про що свідчать високі показники косинусної подібності ($\sim 0.87 \pm 0.03$ у разі моделі векторизації `xx_ent_wiki_sm`). Однак часто, хоча і не у всіх випадках, створена фраза може виглядати дещо неприродною, містити зайві сутності та вирази. Отримані результати вказують на те, що, хоча великі мовні моделі можна застосовувати для створення тексту на основі онтологічного представлення та передачі загального значення, згенеровані фрази часто не ідеальні з огляду на форму (а іноді ще й - нюансів значення). Запропонований у роботі підхід вбачається перспективним в аспекті надання користувачеві природномовних відповідей на основі результатів виконання запитів до баз знань онтологічного типу.

Подяки

Дослідження виконано за підтримки гранту НФДУ (Open Funder Registry: 10.13039/100018227) за проектом

№ 2021.01/0136 (2022–2024, проєкт у процесі реалізації) “Розробка хмарної платформи пацієнт-центричної телереабілітації онкологічних хворих на основі математичного моделювання” [6] на базі Інституту кібернетики імені В.М. Глушкова НАН України.

Література

1. K. Malakhov, M. Petrenko, E. Cohn, Developing an ontology-based system for semantic processing of scientific digital libraries, *South African Computer Journal*, 2023. Vol. 35, No. 1. P. 19–36.
2. O. Palagin, M. Petrenko, M. Boyko, Ontology-related Complex for Semantic Processing of Scientific Data. Proceedings of the 13th International Scientific and Practical Programming Conference UkrPROG 2022. Kyiv, Ukraine, October 11–12, 2022. Vol. 3501. P. 279 – 290.
3. M. Petrenko, E. Cohn, O. Shchurov, K. Malakhov, Ontology-Driven Computer Systems: Elementary Senses in Domain Knowledge Processing. *South African Computer Journal*, 2023. Vol. 35, No.2. P. 127 – 144.
4. K. S. Malakhov, Insight into the Digital Health System of Ukraine (eHealth): Trends, Definitions, Standards, and Legislative Revisions. *International Journal of Telerehabilitation*, 2023. Vol. 15, No. 2. P. 1 – 21.
5. K. S. Malakhov, Letter to the Editor – Update from Ukraine: Development of the Cloud-based Platform for Patient-centered Telerehabilitation of Oncology Patients with Mathematical-related Modeling. *International Journal of Telerehabilitation*, 2023. Vol. 15, No. 1. P. 1–3.
6. K. Malakhov, Letter to the Editor – Update from Ukraine: Rehabilitation and Research. *International Journal of Telerehabilitation*, 2022. Vol. 14, No. 2. P. 1–2.
7. H. Inefuku, K. Malakhov, E. R. Cohn, L. B. Collister, Service Diversification, Connections, and Flexibility in Library Publishing: Rapid Publication of Research from Ukraine in Wartime. *Case Studies in Library Publishing*, 2023. Vol.1, No.1.
8. O. V. Palagin, K. S. Malakhov, V. Yu. Velychko, T. V. Semykopr, Hybrid e-rehabilitation services: SMART-system for remote support of rehabilitation activities and services. *International Journal of Telerehabilitation*, Special Issue: Research Status Report – Ukraine, 2022. P. 1–16.
9. O. Palagin, V. Kaverinskiy, K. Malakhov, M. Petrenko, Fundamentals of the Integrated Use of Neural Network and Ontolinguistic Paradigms: A Comprehensive Approach. *Cybern. Syst. Anal.*, 2024. Vol. 60. P. 111–123.
10. A. A. Litvin, V. Yu. Velychko, V. V. Kaverinsky, Synthesis of chat-bot responses in the natural language of the flexive type based on the results of

formal questions to ontology and semantic analysis of the initial phrase. *International Journal "Information Content and Processing"*, 2020. Vol. 7, No. 1.

11. O. Palagin, V. Kaverinskiy, K. Malakhov, A. Litvin, OntoChatGPT Information System: Ontology-Driven Structured Prompts for ChatGPT Meta-Learning. *International Journal of Computing*, 2023. Vol. 22, No. 2. P. 170 – 183.
12. Levenshtein distance. [cited 20 Mar 2024]. https://en.wikipedia.org/wiki/Levenshtein_distance
13. A.F. Kurgaev, N.G. Petrenko, Processor structure design. *Cybern Syst Anal.*, 1995. Vol. 31. P. 618–625.
14. N. G. Petrenko, A. A. Sofiyuk, On one approach to the transfer of an information structures interpreter to PLD-implementation. *Upravlyayushchie Sistemy i Mashiny*, 2003. No. 6. P. 48 – 57.

Одержано: 21.03.2024

Внутрішня рецензія отримана: 14.04.2024

Зовнішня рецензія отримана: 20.04.2024

Про авторів:

¹*Каверинський Владислав Володимирович*,
кандидат технічних наук,
старший дослідник,
старший науковий співробітник.
<https://orcid.org/0000-0002-6940-579X>.

¹*Літвін Анна Андріївна*,
кандидат технічних наук,
науковий співробітник.
<http://orcid.org/0000-0002-5648-9074>.

¹*Палагін Олександр Васильович*,
доктор технічних наук,
академік НАН України,
заступник директора
Інституту кібернетики НАН України.
<https://orcid.org/0000-0003-3223-1391>

Місце роботи авторів:

¹Інститут кібернетики НАН України,
Тел. (+38) (044) 526-20-08
E-mail: incyb@incyb.kiev.ua,
www.incyb.kiev.ua