



Державна наукова установа
«Енциклопедичне видавництво»

**УКРАЇНСЬКА
ЕНЦИКЛОПЕДИСТИКА
В УМОВАХ
ВОЄННОГО СТАНУ.
ПИТАННЯ ТЕОРІЇ
ТА ПРАКТИКИ**



ДЕРЖАВНА НАУКОВА УСТАНОВА
«ЕНЦИКЛОПЕДИЧНЕ ВИДАВНИЦТВО»

**УКРАЇНСЬКА ЕНЦИКЛОПЕДИСТИКА
В УМОВАХ ВОЄННОГО СТАНУ.
ПИТАННЯ ТЕОРІЇ ТА ПРАКТИКИ**



Київ 2024

УДК 655.3.066.11:031]”364”(477)(0.064)
ББК Ч617.167.7(4УКР),021
У45

*Рекомендовано до друку Вченою радою
Державної наукової установи «Енциклопедичне видавництво»
(Протокол № 4 від 28 червня 2024 р.)*

Авторський колектив: д. і. н., проф. А. Киридон (заг. ред., керівник авторського колективу); д. філос. н., проф. А. Арістова; к. арх. В. Вечерський; к. і. н. Н. Гаврилишина; д. тех. н., проф. Г. Гайко; к. тех. н., доц. А. Гладун; к. і. н., доц. Л. Дояр; к. біол. н. Л. Зав'ялова; д. пед. н., проф. Т. Кочубей; к. г. н., доц. Л. Кушнір; к. е. н., доц. Л. Кушнір; д. тех. н., проф. Д. Ланде; М. Моторна; к. філос. н., с. н. с. О. Недавня; К. Подолець; к. філол. н. Л. Рева; Л. Самчук; д. і. н., проф. С. Троян; К. Хала; к. і. н. М. Чорна; В. Яремчук.

Рецензенти:

*Дорошко М. С., доктор історичних наук, професор
Мачуський Є. А., доктор історичних наук, професор
Титаренко В. В., доктор філософських наук, професор*

Українська енциклопедистика в умовах воєнного стану. Питання теорії та практики: кол. монографія / Заг. ред. д. і. н., проф. Алли Киридон.
У45 Київ : Державна наукова установа «Енциклопедичне видавництво», 2024. 152 с.

ISBN 978-617-8272-07-4

Реальність, породжена від 2014 р. російсько-українською війною, а особливо її кульмінаційною фазою — широкомасштабним вторгненням Росії в Україну, визначила не тільки нову екзистенційну парадигмальність, але й обумовила необхідність переосмислення складного комплексу проблем в просторі енциклопедистики (перегляд усталених принципів, матриць і критеріїв формування контенту, інформаційно-методологічних парадигм, осмислення подієвої історії тощо). Пропонована праця колективу авторів репрезентує широкий спектр дослідницьких рефлексій в царинах теорії, методології, історії енциклопедистики, формування стилю та контенту, інформаційного забезпечення та функціонування сучасних енциклопедій тощо.

Видання розраховане на науковців, викладачів, студентів, працівників видавництва довідкової літератури та всіх, хто цікавиться питаннями енциклопедистики й становленням знанневої парадигми в енциклопедійному вимірі.

ISBN 978-617-8272-07-4

©Державна наукова установа
«Енциклопедичне видавництво», 2024

2.5. Критерії добору персоналій-представників української діаспори до «Великої української енциклопедії» (<i>Мар'яна Чорна</i>)	73
2.6. Подієва історія в енциклопедистиці: топографія останнього рейду Нестора Махна (13–28.08.1921) (<i>Леонід Кушнір, Людмила Кушнір</i>).....	79

РОЗДІЛ III КУЛЬТУРНІ ПРАКТИКИ В ЕНЦИКЛОПЕДИЧНОМУ ДИСКУРСІ

3.1. Репрезентація релігійно-звичаєвої спадщини старокиївської культурної традиції в енциклопедійному просторі (<i>Ольга Недавнія</i>).....	88
3.2. Репрезентація у «Великій українській енциклопедії» пам'яток всесвітньої спадщини ЮНЕСКО (<i>Віктор Вечерський</i>).....	92
3.3. Досвід збереження нематеріальної культурної спадщини в умовах сьогодення (<i>Тетяна Кочубей, Марія Моторна</i>)	100
3.4. Бібліотека — необхідна частина соціокультурної інфраструктури суспільства (<i>Лариса Рева</i>)	104
3.5. Перший випуск науково-допоміжного бібліографічного покажчика «Медичне забезпечення антитерористичної операції та операції об'єднаних сил» (<i>Лариса Самчук</i>)	108

РОЗДІЛ IV ВИКОРИСТАННЯ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ В ЕНЦИКЛОПЕДИСТИЦІ

4.1. Енциклопедичний дискурс комунікаційних платформ: від мережевих енциклопедій — до енциклопедійної мережі (<i>Алла Арістова</i>)	114
4.2. Особливості застосування енциклопедійних ресурсів для розширення онтологічних моделей як засобу керування взаємодією дронів (<i>Анатолій Гладун, Катерина Хала</i>)	125
4.3. Формування та використання семантичних мереж для навігації в енциклопедичних даних (<i>Дмитро Ланде</i>)	142
Відомості про авторів.....	147

4.3. ФОРМУВАННЯ ТА ВИКОРИСТАННЯ СЕМАНТИЧНИХ МЕРЕЖ ДЛЯ НАВІГАЦІЇ В ЕНЦИКЛОПЕДИЧНИХ ДАНИХ

Сучасні енциклопедії, що представлені в Інтернеті, по суті є семантичними мережами, оскільки їх статті зв'язані гіперпосиланнями. Як правило, кожна стаття в енциклопедії не існує ізольовано, а пов'язана з іншими через гіперпосилання, що створює складну мережу знань. Це дає змогу користувачам швидко переходити від однієї теми до іншої, що робить навігацію інформацією більш ефективною та зручною.

Проте сучасні технології, такі як великі лінгвістичні моделі, або LLMs (від англ. — Large Language Model)²⁴ відкривають нові можливості для покращення семантичного індексування енциклопедій. Відомі великі лінгвістичні моделі, які можуть бути застосовані до цього процесу, включають, наприклад, GPT-4 від OpenAI²⁵, BERT від Google²⁶, Meta Llama²⁷. Ці моделі здатні аналізувати великі обсяги тексту, розуміти контекст і знаходити смислові зв'язки між різними частинами інформації.

Використання таких моделей для семантичного індексування, зокрема, «Великої української енциклопедії» дозволить виявляти додаткові вузли-поняття та зв'язки між ними. Семантичне індексування за допомогою LLMs дозволить не лише створити більш точну і детальну карту знань, але й знайти нові, раніше невідомі зв'язки між поняттями-статтями, що може суттєво покращити якість і зручність навігації в енциклопедичних даних.

У результаті семантичного індексування можливо формування розширених семантичних мереж, які зможуть бути використані для навігації в енциклопедичних даних. Користувачі отримують можливість не лише швидко знаходити необхідну інформацію, але й розуміти її контекст, що зробить процес навчання і пошуку ефективнішим.

Крім того, розширені семантичні мережі можуть бути використані для постановки нових задач у розвитку енциклопедії, додавання нових понять, встановлення нових зв'язків. Завдяки цьому, енциклопедія може

²⁴ Lande D., Strashnoy L. GPT Semantic Networking: A Dream of the Semantic Web — The Time is Now. Kyiv : Engineering, 2023. 168 p.

²⁵ GPT-4 is OpenAI's Most Advanced System, Producing Safer and More Useful Responses // OpenAI. URL: <https://openai.com/index/gpt-4/>

²⁶ Google-Research / Bert // Github. URL: <https://github.com/google-research/bert>

²⁷ Meet Llama 3.1 // Meta. URL: <https://llama.meta.com/>

постійно оновлюватися і вдосконалюватися, ставати більш актуальною і корисною для користувачів. Тим самим великі лінгвістичні моделі допоможуть не лише покращити навігацію і використання даних, але й сприяють їх постійному розвитку та збагаченню енциклопедії.

Мета роботи полягає у створенні ефективного інструменту для покращення навігації, пошуку і використання енциклопедичних даних, а також для постійного розвитку і збагачення енциклопедії за допомогою сучасних технологій і методів семантичного індексування. Цим самим мають бути вирішені задачі: покращення навігації та пошуку інформації; виявлення нових знань та зв'язків; покращення якості енциклопедичних даних; забезпечення контекстуальної розумності; розвитку та збагачення енциклопедії.

Методологія семантичного індексування документів

Семантичне індексування документів є ключовим етапом у створенні розширених семантичних мереж, які забезпечують зручну навігацію та пошук інформації в енциклопедичних даних. Запропонована методологія цього процесу включає кілька послідовних етапів.

Першим кроком є визначення та вибір статей, які належать до певної предметної галузі. Це дозволяє сфокусувати зусилля на обмеженому колі тем, що забезпечує глибше розуміння та аналіз інформації. Наприклад, для галузі кібербезпеки можуть бути вибрані статті, що охоплюють теми криптографії, мережевої безпеки, захисту інформації тощо.

Другим кроком використовуючи великі лінгвістичні моделі, зокрема, GPT-4, здійснюється витягування пар зв'язаних понять із вибраних текстів. Це досягається за допомогою спеціально сформульованих промптів, які спрямовують модель на виявлення зв'язків між різними брендмауерами, тощо.

Третій крок — це формування семантичної мережі шляхом накопичення пар зв'язаних понять. Кожна пара понять додається як вузол і ребро в цю мережу, що дозволяє створити структуровану карту знань, яка відображає взаємозв'язки між різними поняттями в межах певної галузі.

Четвертий крок — візуалізація мережі за допомогою створеного автором сервісу CSV2Graph²⁸. Цей інструмент приймає на вхід дані у форматі CSV, що містять пари зв'язаних понять, і генерує візуальні карти мереж у

²⁸ Semantic Networking // Dmytro Lande. URL: <https://bigsearch.space/uli.html>

форматі SVG. В процесі візуалізації сервіс також додає посилання на різні пошукові бази даних, включаючи мережеву енциклопедію Wikipedia²⁹, що дозволяє користувачам легко переходити до додаткових ресурсів для отримання більш детальної інформації.

Приклад із сфери кібербезпеки

Для ілюстрації методології розглянемо приклад із галузі кібербезпеки. Розглядалися статті за темами «Кіберзлочин», «Кіберзброя», «Комп'ютерне шахрайство» тощо. Для цього застосовувались промпти до LLM GPT-4 на кшталт:

→Виведіть із тексту 20 пар найбільш пов'язаних сутностей. Кожна сутність має бути описана не більше ніж 3 словами. Формат виведення: “сутність 1; сутність 2”. Кожна пара відображається в новому рядку. Текст: Комп'ютерне шахрайство (кібершахрайство) — це виведення, заміна, виправлення, знищення комп'ютерних даних або програм чи інші втручання до процесу обробки інформації, які впливають на кінцевий результат, спричиняють економічні або майнові збитки з метою одержання незаконного економічного прибутку для себе чи іншої особи...

На Рис. 1 як приклад показано результат — фрагмент мережі предметної області «Кібербезпека», який було сформовано на базі вибраних статей.

Найвагомішими вузлами побудованої семантичної мережі виявилися такі сутності: «Cybercrime», «Cyberwarfare», «Melissa (computer virus)», «Computer fraud», «Hacker motivation» (*немає в енциклопедії*), «White hat hacker».

Запропоновані підходи дозволяють вирішити поставлені завдання щодо розвитку енциклопедичного ресурсу, а саме, забезпечать:

- Оптимізацію процесів навігації та пошуку інформації завдяки тому, що розширені семантичні мережі дозволяють створити більш детальну і точну карту знань, полегшать пошук і доступ до необхідної інформації, користувачі можуть швидко переходити між пов'язаними поняттями.
- Виявлення нових знань та зв'язків завдяки семантичному індексуванню та аналізу за допомогою LLMs, які дозволяють виявляти нові, неочевидні зв'язки між поняттями.

²⁹ Wikipedia (офіційний сайт). URL: <https://en.wikipedia.org/>

- Покращення якості енциклопедичних даних шляхом семантичного індексування, яке допомагає знаходити і виправляти помилки, дублікати та інші неточності в енциклопедичних даних.

- Розвиток та збагачення енциклопедії за рахунок ідентифікації прогалини в наявних даних та додання нових понять і зв'язків, що забезпечує постійне оновлення та розвиток енциклопедії.

Крім того, семантичні мережі допомагають користувачам розуміти контекст інформації, яку вони знаходять, що необхідно для глибшого розуміння матеріалу і його правильного застосування.