

DOI 10.33930/ed.2019.5007.52(1-2)-13
УДК 161:004.8

ТЕСТУВАННЯ СЕРВІСУ ШТУЧНОГО ІНТЕЛЕКТУ DEEPAI.ORG НА ЗДАТНІСТЬ ДОТРИМУВАТИСЯ ПРАВИЛ ФОРМАЛЬНОЇ ЛОГІКИ

В. Д. Піхорович
А. Ю. Самарський

Анотація.

Актуальність теми дослідження. Швидкий прогрес інформаційних технологій протягом останніх десятиліть істотно вплинув на багато аспектів суспільства. Особливу увагу привертає поширення використання штучного інтелекту (ШІ). Він стає все більш актуальним і широко застосовується в різних сферах науки, освіти і техніки. Це ставить перед філософами завдання провести ґрунтовні дослідження ШІ та розробити пропозиції стосовно його використання.

Постановка проблеми. Одним з ключових питань, яке виникає при розгляді можливостей ШІ, є його здатність дотримуватися правил формальної логіки. Формальна логіка є основою раціонального мислення і використовується для формалізації процесів мислення та доведення правил та законів. Тому, важливо перевірити, наскільки ШІ може виконувати завдання, які вимагають логічного мислення.

Аналіз останніх досліджень і публікацій. Дослідження великих мовних моделей штучного інтелекту (LLM) на дотримання ними правил формальної логіки проводилися колективами вчених під керівництвом: Р. Нішантома, Й. Вана, М. Амірізаніані, Т. Лью, Х. Чі. Роботи С. Ванга демонструють, що оволодіння основними правилами умовиводів ШІ все ще не відповідає людським можливостям та пропонують підходи для перевірки та підвищення логічних навиків ШІ. Е. Ларсон, показує, чому ШІ не може розуміти тексти, проте він спеціально не досліджує дотримання ШІ правил формальної логіки. Б. Лін досліджує використання математичних методів для підвищення точності ШІ.

Дослідження показують, що ШІ може бути навченим дотримуватися правил формальної логіки, але цей процес може вимагати значних зусиль та ресурсів.

Постановка завдання. Завдання статті полягає в тестуванні сервісу штучного інтелекту <https://deepai.org/chat> на здатність дотримуватися правил формальної логіки, а саме розв'язувати задачі на *modus ponens* і *modus tollens*. Розуміння цього аспекту дозволить краще зрозуміти можливості мовних ШІ і допоможе їх користувачам отримувати більш адекватні результати.

Виклад основного матеріалу. Стаття присвячена аналізу логічних умовиводів, зокрема правил *modus ponens* і *modus tollens* у контексті взаємодії з штучним інтелектом (ШІ). Розглядається два приклади. У першому випадку з двох засновків ("Якщо на вулиці йде дощ, то асфальт мокрий" і "Асфальт на вулиці мокрий") ШІ робить неправильний висновок, стверджуючи, що "на вулиці йде дощ", попри те, що це не є коректним застосуванням *modus ponens*. Автори підкреслюють, що навіть здоровий глузд вказує на те, що асфальт може бути мокрим з інших причин.

У другому прикладі, що стосується *modus tollens*, ШІ отримує ще один неправильний висновок. Знову ведеться мова про те, що коректний висновок може бути зроблений лише за правильними схемами логічного мислення. Хоча в окремих випадках ШІ досягав успіху в логічних задачах, автор вважає це випадковістю, а не підтвердженням здатності до логічного мислення.

Тестування на виявлення помилок у висновках показує, що навіть коли ШІ формулює висновки неправильно, він може вказати на необхідність додаткової інформації для коректного висновку. Таким чином, стаття підкреслює обмеженість сучасних моделей ШІ в розумінні логіки і необхідність критичної думки з боку користувача при роботі з ними.

Висновки. Для вирішення проблеми логічних помилок у текстах, згенерованих штучним інтелектом (ШІ), пропонується об'єднати зусилля фахівців з логіки та ШІ. Важливо розробити програми, які автоматично перевірятимуть відповідність створених текстів схемам *modus ponens* і *modus tollens*. Схеми цих умовиводів можуть бути сформульовані у вигляді математичних моделей, що полегшує виявлення помилок. Необхідно навчити ШІ ідентифікувати антецедент і консеквент в умовних судженнях. Важливо також навчати студентів основ формальної логіки та залучати фахівців-філологів для розробки інструментів контролю. З огляду на складність і відмінності мов, міждисциплінарний підхід буде вирішальним для успіху таких проектів.

Ключові слова: штучний інтелект, формальна логіка, мислення, мовні моделі, LLM, deepai.org/chat, *modus ponens*, *modus tollens*.

TESTING THE ARTIFICIAL INTELLIGENCE SERVICE DEEPAI.ORG ON ITS ABILITY TO FOLLOW THE RULES OF FORMAL LOGIC

Vasil Pihorovich
Andrii Samarskyi

Abstract.

Urgency of the research. The rapid progress of information technologies over the past few decades has

significantly impacted many aspects of society. Special attention is drawn to the increasing use of artificial intelligence (AI). It is becoming more relevant and widely applied in various fields of science, education and technology, ranging from medicine to manufacturing automation. This poses a challenge for philosophers to conduct in-depth research on artificial intelligence and develop recommendations regarding its use.

Target setting. One of the key questions arising when considering the capabilities of artificial intelligence is its ability to adhere to the rules of formal logic. Formal logic is the foundation of rational thinking and is used to formalize thought processes and prove rules and laws. Therefore, it is essential to test how well artificial intelligence can perform tasks that require logical thinking.

Actual scientific researches and issues analysis. Research on large language models (LLM) compliance with the rules of formal logic has been conducted by various groups of scientists led by R. Nishant, J. Wan, M. Amirizani, T. Liu, and H. Chi. The work of S. Wang demonstrates that AI's mastery of the fundamental rules of inference still falls short of human capabilities and suggests approaches for testing and enhancing AI's logical skills. E. Larson shows why artificial intelligence cannot understand texts; however, he does not specifically investigate artificial intelligence's adherence to the rules of formal logic. B. Lin explores the use of mathematical methods to improve artificial intelligence accuracy.

Research indicates that artificial intelligence can be trained to follow the rules of formal logic, but this process may require significant effort and resources.

The research objective. The task of this article is to test the artificial intelligence service at <https://deepai.org/chat> for its ability to adhere to the rules of formal logic, specifically solving problems related to modus ponens and modus tollens. Understanding this aspect will allow for a better understanding of the capabilities of language-based artificial intelligence and assist users in obtaining more accurate results.

The statement of basic materials. The article is dedicated to analyzing logical inferences, particularly the rules of modus ponens and modus tollens in the context of interaction with artificial intelligence (AI). Two examples are examined. In the first case, from two premises ("If it is raining outside, then the pavement is wet" and "The pavement outside is wet"), artificial intelligence makes an incorrect conclusion, asserting that "it is raining outside," despite this being an incorrect application of modus ponens. The authors emphasize that common sense indicates that the pavement could be wet for other reasons.

In the second example, concerning modus tollens, artificial intelligence reaches yet another incorrect conclusion. Again, it is noted that a correct conclusion can only be drawn from proper logical thinking schemes. Although artificial intelligence has succeeded in some logical tasks, the author considers this a coincidence rather than confirmation of logical thinking ability.

Testing for errors in conclusions shows that even when artificial intelligence formulates conclusions incorrectly, it can indicate the need for additional information to make a correct inference. Thus, the article highlights the limitations of contemporary artificial intelligence models in understanding logic and underscores the necessity for critical thinking on the part of users when interacting with them.

Conclusions. To address the issue of logical errors in texts generated by artificial intelligence (AI), it is proposed to unite the efforts of specialists in logic and AI. It is important to develop programs that will automatically check the conformity of generated texts to the schemas of modus ponens and modus tollens. The schemas of these inferences can be formulated as mathematical models, facilitating the detection of errors. It is also essential to teach AI to identify antecedents and consequents in conditional statements. Moreover, it is vital to educate students on the fundamentals of formal logic and engage linguists in the development of control tools. Given the complexity and differences in languages, an interdisciplinary approach will be crucial for the success of such projects.

Keywords: artificial intelligence, formal logic, reasoning, language models, LLM, deepai.org/chat, modus ponens, modus tollens.

Актуальність теми дослідження. Швидкий прогрес інформаційних технологій протягом останніх десятиліть істотно вплинув на багато аспектів суспільства. Наука і техніка зазнали значних змін. Практично неможливо знайти наукову дисципліну, яка б не піддалася впливу інформатизації та цифровізації. Особливу увагу привертає поширення використання штучного інтелекту (ШІ). Він стає все більш актуальним і широко застосовується в різних сферах науки, освіти та техніки. Це ставить перед філософами завдання провести ґрунтовні дослідження ШІ та розробити пропозиції стосовно його використання в освітньому процесі. Варто зауважити, що навички логічного мислення не є єдиним показником інтелекту ШІ. Штучний інтелект також вимагає розвитку інших аспектів, таких як розпізнавання образів, навчання на основі даних та прийняття рішень.

Постановка проблеми. Одним з ключових питань, яке виникає при розгляді можливостей ШІ, є його здатність дотримуватися правил формальної логіки. Формальна логіка є основою раціонального мислення і використовується для формалізації процесів мислення та доведення правил та законів. Тому, важливо перевірити, наскільки ШІ може виконувати завдання, які вимагають логічного мислення. Тестування здатності ШІ дотримуватися правил формальної логіки є важливим кроком у розвитку інтелектуальних систем. Це дозволить розширити можливості та збільшити впевненість у використанні ШІ в різних галузях, де вимагається логічне мислення і прийняття обґрунтованих рішень.

Аналіз останніх досліджень і публікацій. Дослідження в сфері тестування ШІ на здатність дотримуватися правил формальної логіки спрямовані на розробку алгоритмів та моделей, що дозволяють оцінювати логічні навички інтелектуальних систем. Один з підходів до цього полягає у створенні комп'ютерних програм, які перевіряють, наскільки ШІ виконує логічні операції та доводить правила. Як правило, над такою задачею працює колектив дослідників великих мовних

моделей LLM (від англ. “large language model”) до яких належить популярний Chat GPT.

Оскільки практиці використання LLM приблизно 7 років, а їх широке використання почалось декілька років тому, то досліджень з даної проблематики ще небагато. Проте можна виділити ключові публікації. У 2021 році вийшла книга Е. Ларсона “Міф про штучний інтелект: чому комп’ютери не можуть мислити так, як ми” [3], де автор показує, чому ШІ не може розуміти тексти, проте він спеціально не досліджує дотримання ШІ правил формальної логіки, хоч і стверджує, що LLM, на відміну від людини, занадто формальний. В статті “Формальна раціональність алгоритмів на основі штучного інтелекту та проблема зміщення” Р. Нішантом та іншими авторами [6] досліджується фундаментальна різниця між ШІ та людською раціональністю в осмисленні даних та чому алгоритмам на основі ШІ не вистачає здатності відображати здоровий глузд в інтерпретації даних, що призводить до помилкових рішень.

Й. Ван та інші автори досліджень “LogicAsker: оцінка та вдосконалення здатності логічного мислення великих мовних моделей” [7] та “A & B == B & A: виклик логічних збоїв у великих мовних моделях” [8] також зафіксували проблему дотриманні ШІ правил логіки та запропонували автоматичний підхід, який всебічно оцінює та покращує здатність логічного міркування LLM за набором атомарних навичок міркування на основі пропозиційної та предикатної логіки. М. Амрізаніані та інші автори публікації “Чи можуть великі мовні моделі міркувати як люди? Оцінювання теорії мислення в LLMs для відкритих запитань” [1] порівнюючи показники семантичної схожості та лексичного перекриття між відповідями, створеними людьми та LLM, виявляє явні відмінності в можливостях аргументації, причому навіть найдосконаліші моделі демонструють помітні обмеження.

Дослідження “Логіка думки: введення логіки в контексти для повного міркування у великих мовних моделях” авторів Т. Лью та ін. [5] констатує, що продуктивність LLM у складних логічних завданнях залишається незадовільною. Інше дослідження авторів Х. Чі та ін. “Розкриття причинно-наслідкових міркувань у великих мовних моделях: реальність чи міраж?” [2] робить висновок, що LLM здатні лише до неглибокого рівня причинно-наслідкових міркувань, але їм не вистачає здатності до справжнього людського рівня причинно-наслідкових міркувань. Автори дослідження “Чи можуть LLM міркувати за правилами? Логічні каркаси для стрес-тестування та покращення LLM” С. Ванг та ін. [9] демонструють, що опанування основними правилами умовиводів ШІ все ще не відповідає людським можливостям та пропонують підходи для перевірки та підвищення логічних навичок ШІ. Крім того, такі компанії, як Amazon, інвестують у методи автоматизованого міркування для зменшення помилок ШІ і забезпечення його відповідності формальної логіки. В статті Б. Лін “Чому Amazon робить ставку на “автоматизоване міркування”, щоб зменшити галюцинації ШІ?” [4] обговорюється використання математичних методів для підвищення точності ШІ. Дослідження показують, що ШІ може бути навченим дотримуватися правил формальної логіки, але цей процес вимагатиме значних зусиль та ресурсів.

Постановка завдання. Завдання статті полягає в тестуванні сервісу штучного інтелекту <https://deepai.org/chat> на здатність дотримуватися правил формальної логіки, а саме розв’язувати задачі на *modus ponens* і *modus tollens*. Розуміння цього аспекту дозволить краще зрозуміти можливості мовних ШІ і допоможе їх користувачам отримувати більш адекватні результати.

Виклад основного матеріалу. Поставимо задачу на *modus ponens* (так на латині називається стверджувальний модус умовно-категоричного умовиводу):

“Чи можна зробити висновок “На вулиці йде дощ” із засновків “Якщо на вулиці йде дощ, то асфальт мокрий” і “Асфальт на вулиці мокрий””. Відповідь чату: “Так, можна зробити висновок “На вулиці йде дощ”, оскільки з двох засновків можна вивести, що якщо на вулиці йде дощ, то асфальт стає мокрим, а асфальт на вулиці дійсно мокрий. Таким чином, можна зробити висновок про те, що на вулиці йде дощ”.

Відповідь очевидно некоректна, оскільки ствердження за схемою *modus ponens* може йти тільки від ствердження основи до ствердження наслідку, але не навпаки — від ствердження наслідку до ствердження основи, оскільки одне і те ж явище може викликатися різними причинами, і наявність цього явища зовсім не значить, що воно викликане саме тією причиною, про яку йшлося у засновку.

У нашому випадку основою служило твердження “на вулиці йде дощ” (позначимо його буквою А), а наслідок був “асфальт мокрий” (його позначимо буквою В). Отже:

Правильна схема

$$\begin{array}{l} A \Rightarrow B \\ A \\ \hline B \end{array}$$

Неправильна схема

$$\begin{array}{l} A \Rightarrow B \\ B \\ \hline A \end{array}$$

Для того, щоб розуміти, що такого висновку, як зробив у цьому випадку штучний інтелект,

робити не можна, зовсім не обов'язково знати правила побудови умовиводів за схемою *modus ponens*. Досить звичайного здорового глузду, щоб побачити, що асфальт міг бути мокрий, скажімо, від того, що розтав сніг, чи його полила поливальна машина.

Спробуємо ще одну задачу, але на заперечення, тобто на міркування за схемою *modus tollens*, заперечний модус умовно-категоричного умовиводу. Сформулюємо запит так: “Чи можна зробити висновок “Асфальт на вулиці не мокрий” із засновків “Якщо на вулиці йде дощ, то асфальт мокрий” і “На вулиці не йде дощ””.

На цей раз результат вийшов ще гірший, оскільки на додаток до порушення правила *modus tollens*, згідно з яким заперечення може йти тільки від заперечення наслідку до заперечення основи, у першому реченні ми отримали чисту нісенітницю:

“З висновком “Асфальт на вулиці не мокрий” можна зробити наступний логічний висновок на основі двох даних: “Якщо на вулиці не йде дощ, то асфальт на вулиці не мокрий”. Це впливає з умов: якщо на вулиці йде дощ, то асфальт мокрий, але оскільки на вулиці не йде дощ, то асфальт не мокрий”.

Як уже говорилося, заперечення за схемою *modus tollens* може здійснюватися тільки від заперечення наслідку до заперечення основи. Причина такого порядку заперечення та ж сама, що і у випадку зі ствердженням: будь-яке явище може викликатися різними причинами, отже відсутність тієї, про яку йдеться у засновку не означає, що явище не може бути викликане іншими причинами. А от відсутність явища означає, що не було ніяких причин, у тому числі та тієї, яка нас цікавить. Таким чином:

Правильна схема

$$\begin{array}{l} A \Rightarrow B \\ \text{не } B \\ \hline \text{не } A \end{array}$$

Неправильна схема

$$\begin{array}{l} A \Rightarrow B \\ \text{не } A \\ \hline \text{не } B \end{array}$$

Виходить, що для даного сервісу ІІІ цей алгоритм не є очевидним.

Але нездатність вирішувати навіть найпростіші логічні задачі на *modus ponens* чи *modus tollens* ще не означає, що він взагалі не здатний будувати міркування за цими схемами. Адже нездатність вирішити логічну задачу для ІІІ може бути пов'язана зі способом, яким ця задача була сформульована. Адже ІІІ — це не інтелект у звичному розумінні цього слова, а мовна модель. Він думати не вміє, а може тільки продовжувати фразу тим чи іншим способом. Тому поставимо перед чатом задачу трохи інакше: “Відомо, що коли на вулиці йде дощ, то асфальт мокрий. Зараз на вулиці йде дощ. Який висновок можна зробити із цих даних?”

Сервіс видав цілком коректну відповідь: “З цих даних можна зробити висновок, що асфальт на вулиці мокрий”. Зрозуміло, що робити на основі цієї відповіді висновок про те, що програми штучного інтелекту можуть бути навчені розв'язувати задачі логічного мислення, було б не дуже коректно. Коректніше було б говорити про те, що в нескладних випадках, коли засновки ідуть за правильною схемою *modus ponens*, він здатен зробити коректний висновок.

Перевіримо, чи зможе сервіс повторити своє “досягнення” на прикладі з *modus tollens*. Сформулюємо запит: “Відомо, що коли на вулиці йде дощ, то асфальт мокрий. Зараз асфальт не мокрий. Який висновок можна зробити із цих даних?” Відповідь правильна: “З цих даних можна зробити висновок, що зараз не йде дощ на вулиці”.

Отже, і за схемою *modus tollens* у нескладному випадку, тобто за умови коректного формулювання засновків, сервіс може зробити адекватний висновок. Але чи можна це характеризувати як вміння розв'язувати задачі логічного мислення. Адже може статися, що коректні висновки з'являються цілком випадково.

Тому поспробуємо визначити межі складності логічних задач на тему умовно-категоричних умовиводів, доступної даному сервісу. Можливо, що ІІІ не бачив логічних помилок тільки тому, там їх треба було шукати в уже готових умовиводах, але зможе вичленити їх тоді, коли йому самому доведеться робити висновок. Сформулюємо запит так, щоб засновки, із яких ІІІ має зробити висновки, були істинними, але йшли за неправильною схемою *modus tollens*:

$$\begin{array}{l} A \Rightarrow B \\ \text{не } A \\ \hline \text{не } B \end{array}$$

“Відомо, що коли на вулиці йде дощ, то асфальт мокрий. Зараз на вулиці не йде дощ. Який висновок можна зробити із цих даних?” Знову отримуємо помилкову відповідь: “З цих даних можна зробити висновок, що асфальт не є мокрим”. Перевіримо, що буде, коли засновки будуть побудовані за неправильним *modus ponens*:

$$\frac{A \Rightarrow B}{B} \\ \frac{\quad}{A}$$

“Відомо, що коли на вулиці йде дощ, то асфальт мокрий. Зараз асфальт мокрий. Який висновок можна зробити із цих даних?” Відповідь знову некоректна: “З цих даних можна зробити висновок, що асфальт на вулиці зараз не мокрий, оскільки не йде дощ”.

Отже, даний сервіс не в змозі не тільки виявити логічну помилку в готових міркуваннях, побудованих за схемами *modus ponens* чи *modus tollens*, але і не здатен виявляти випадки, коли висновку не можна робити через те, що засновки задані з порушенням правил.

Проте можна “змусити” його знайти помилку. Зробимо наступний запит:

“Відомо, що коли на вулиці йде дощ, то асфальт мокрий. Зараз асфальт мокрий. Який висновок можна зробити із цих даних?”

Некоректна відповідь:

“З цих даних можна зробити висновок, що асфальт на вулиці зараз не мокрий, оскільки не йде дощ”.

Адже асфальт може бути мокрим з інших причин”.

Отримуємо наступну відповідь сервісу: “наприклад, через те, що раніше йшов дощ або що було пролите якоюсь рідиною. Тому правильний висновок з цих даних може бути такий: “Немає достатніх даних для висновку щодо того, чому асфальт зараз мокрий””.

Виходить, що в принципі ШІ може робити цілком коректний висновок про те, що коли в засновках містяться логічні помилки, то ми не маємо права робити висновки з них. Але для того, щоб він зробив цей висновок, потрібні додаткові зусилля користувача. До того, щоб ці зусилля докласти, користувач сам має або знати правила побудови *modus ponens* чи *modus tollens*, або добре розуміти предмет, про який йдеться, бо інакше він просто не зможе побачити тих помилок, які допустив ШІ і, відповідно, ніяких додаткових запитів створювати не буде.

Висновки. Можливо, задля хоча б часткового вирішення цих проблем фахівцям з логіки і ШІ треба об'єднати свої зусилля для того, щоб розробити програми, які б автоматично перевіряли згенеровані ШІ тексти на їх відповідність правилам побудови *modus ponens* і *modus tollens*. Адже, самі схеми умовно-категоричних умовиводів є дуже простими і, що дуже важливо, існують не тільки чіткі алгоритми побудови правильних міркувань за *modus ponens* і *modus tollens*, але не менш чіткі схеми неправильних умовно-категоричних умовиводів, що полегшує задачу створення програм для виявлення характерних помилок при генеруванні текстів.

Ці схеми являють собою готові математичні моделі для будь-яких умовно-категоричних умовиводів, тобто таких умовиводів, першим засновком яких служать умовні, інакше кажучи, імплікативні судження, а другим — прості категоричні судження, до того ж такі прості судження, які входять до складу умовного судження першого засновку як його антецедент (умова) чи наслідок (консеквент). Фактично головна проблема при розробці програм, здатних виявляти і виправляти логічні помилки в міркуваннях, побудованих за такими схемами, може полягати в тому, щоб навчити ШІ ідентифікувати в умовних судженнях, які містяться у першому засновку умовно-категоричних умовиводів, умову і наслідок, або інакше — антецедент і консеквент.

Дослідження показало, що потрібно обов'язково вчити основ формальної логіки студентів тих спеціальностей, які в майбутньому збираються займатися розробкою ШІ, і тих, хто змушений буде фахово використовувати ШІ в своїй роботі. Для вирішення цієї проблеми було б доречно залучити до роботи і фахівців-філологів. Адже ознаки, за якими можлива така ідентифікація, далеко не завжди зводяться до тих, які відомі нам зі шкільної програми. Наприклад, що слова, які будуть іти після “якщо” (у нашому прикладі це “йде дощ”) складають умову, а ті, які слідує за “то”, наслідок (у нашому прикладі це “йде дощ”). Зрозуміло, що ШІ буде працювати не тільки з такими очевидними граматичними конструкціями, які були використані у нашому прикладі, не кажучи вже про те, що зміст речень далеко не завжди буде таким же очевидним, що дозволяє користувачу самому легко контролювати ШІ, якщо він припуститься помилку. Крім того, треба мати на увазі, що “робочою мовою” ШІ зазвичай є англійська, тому у випадку з іншими мовами ми будемо мати справу ще із подвійним перекладом, який теж не буде сприяти коректній ідентифікації антецедента і консеквента. Тому роль лінгвістів при розробці інструментів логічного контролю ШІ буде дуже велика. А якщо взяти до уваги, причинно-наслідкові зв'язки, виражені у першому засновку умовно-категоричних суджень далеко не завжди такі очевидні, як у прикладі з дощем і асфальтом, то цілком може бути, що доведеться залучити також і фахівців найрізноманітніших галузей науки і спочатку створювати *modus ponens* і *modus tollens* спеціалізовані інструменти логічного контролю ШІ і тільки потім поступово рухатися шляхом їх універсалізації. На даній стадії дослідження цього питання це питання поки що залишається відкритим, але і тепер зрозуміло, що створення подібних інструментів принаймні на

перших порах може бути тільки результатом міждисциплінарного підходу.

References:

1. Amirizani, M, Martin E, Sivachenko, M, Mashhadi, A & Shah, C 2024, 'Can LLMs reason like humans? Assessing theory of mind reasoning in LLMs for open-ended questions', in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pp. 34-44.
2. Chi, H, Li, H, Yang, W, Liu, F, Lan, L, Ren, X & Han, B 2024, 'Unveiling causal reasoning in large language models: reality or mirage?', in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*. Available from : <<https://openreview.net/pdf?id=1IU3P8VDbn>>. [5 January 2025].
3. Larson, EJ 2021, *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*. Cambridge, MA: Harvard University Press.
4. Lin, B 2025, 'Why Amazon is betting on 'automated reasoning' to reduce AI's hallucinations', *The Wall Street Journal*, 5 February. Available from : <<https://www.wsj.com/articles/why-amazon-is-betting-on-automated-reasoning-to-reduce-ais-hallucinations-b838849e>>. [7 January 2025].
5. Liu, T, Xu, W, Huang, W, Wang, X, Wang, J, Yang, H & Li, J 2024, 'Logic-of-thought: injecting logic into contexts for full reasoning in large language models', *arXiv preprint arXiv:2409.17539*. Available from : <<https://doi.org/10.48550/arXiv.2409.17539>>. [2 January 2025].
6. Nishant, R, Schneckenberg, D & Ravishankar, MN 2024, 'The formal rationality of artificial intelligence-based algorithms and the problem of bias', *Journal of Information Technology*, 39(1), pp. 19–40. Available from : <<https://doi.org/10.1177/026839622311768>>. [6 January 2025].
7. Wan, Y, Wang, W, Yang, Y, Yuan, Y, Huang, JT, He, P & Lyu, M 2024, 'LogicAsker: evaluating and improving the logical reasoning ability of large language models', in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 2124-2155. Available from : <<https://doi.org/10.48550/arXiv.2401.00757>>. [9 January 2025].
8. Wan, Y, Wang, W, Yang, Y, Yuan, Y, Huang, JT, He, P & Lyu, MR 2024, 'A & b== b & a: triggering logical reasoning failures in large language models', *arXiv preprint arXiv:2401.00757*. Available from : <<https://arxiv.org/abs/2401.00757v1>>. [9 January 2025].
9. Wang, S, Wei, Z, Choi, Y & Ren, X 2024, 'Can LLMs reason with rules? Logic scaffolding for stress-testing and improving LLMs', *arXiv preprint arXiv:2402.11442*. Available from : <<https://doi.org/10.48550/arXiv.2402.11442>>. [8 January 2025].

Прорецензовано / Reviewed

21.01.2025

Рекомендовано до друку / Recommended for publication

11.02.2025

DOI 10.33930/ed.2019.5007.52(1-2)-14

УДК 37.09:004.8:316.774

СТРУКТУРА КОМПЕТЕНТНОСТЕЙ ОСВІТЯН У ГАЛУЗІ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ЦИФРОВОМУ СУСПІЛЬСТВІ

М. А. Умрик

Анотація.

Актуальність теми дослідження. Сфера освіти останні декілька років знаходиться під впливом нових потужних викликів, реагувати на які намагаються науковці, освітяни, дослідники багатьох країн. Одним з таких викликів стала епоха ШІ, що пронизала собою усе цифрове суспільство і освітню галузь зокрема.

Постановка проблеми. Отже, постала потреба розвитку компетентностей у галузі використання ШІ у освітан і необхідність створення відповідної структури компетентностей у галузі ШІ для освітан.

Аналіз останніх досліджень і публікацій. Українські вчені активно досліджують можливості застосування цифрових технологій, в тому числі штучного інтелекту в освітньому процесі. Це такі науковці як: В. Биков, М. Жалдак, О. Кузьмінська, В. Кухаренко, В. Лапінський, М. Олійник, Н. Морзе, Ю. Рамський, С. Семеріков, Є. Смирнова-Трибульська, О. Спірін, О. Струтинська, А. Стрюк, Ю. Триус, В. Франчук та інші.

Постановка завдання. Проаналізувати наукові джерела та релевантні публікації за темою дослідження. Визначити складові компетентностей у галузі ШІ для освітан в цифровому суспільстві. Визначити структуру компетентностей у галузі ШІ для освітан. Окреслити напрями подальших досліджень у галузі використання ШІ освітянами.

Виклад основного матеріалу. Авторами було розроблено трьохрівневі складові компетентностей у галузі ШІ для освітан в цифровому суспільстві, які включають загальні компетентності 21 століття, цифрові компетентності, та специфічні професійні компетентності у галузі ШІ для освітан, а також