

*Домаренко М. В.,**викладач кафедри германської філології та перекладу**Національного університету "Полтавська політехніка імені Юрія Кондратюка"*

ВИКОРИСТАННЯ ПРОГРАМ НА ОСНОВІ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ ТИПУ ТЕКСТ-В-ЗОБРАЖЕННЯ ДЛЯ ПРОВЕДЕННЯ ЗАНЯТЬ З АНГЛІЙСЬКОЇ МОВИ В ЗВО

Анотація. В статті надається стисле пояснення сутності генеративного штучного інтелекту (ГШІ) в цілому і ГШІ типу текст-в-зображення зокрема та проводиться огляд останніх досліджень з його застосування у сфері освіти.

ГШІ – є явищем новим та складним, а його застосування може змінюватись в залежності від сфери та типу самого ГШІ. По цій причині стаття присвячена окремій частині цього глобального питання, а саме аналізу використання програм на основі ГШІ типу текст-в-зображення для проведення занять з англійської мови в ЗВО.

Оскільки ГШІ вийшов на публічний простір лише недавно, його застосування в сфері освіти лише починає відбуватись.

Основний спосіб застосування програм на основі ГШІ типу текст-в-зображення для проведення занять з англійської мови в ЗВО – це створення ілюстрацій та різних візуальних матеріалів, що може допомогти у вивченні нової лексики та візуалізації складних понять та концепцій.

Задля ефективного використання зазначених програм викладачі повинні оволодіти навичками створення точних та правильних запитів, які допоможуть отримати релевантний результат.

Для виконання поставленого завдання ГШІ необхідні точно надані інструкції. Чим точнішими і наповненими необхідною інформацією будуть підказки у запиті, тим менше ГШІ буде додумувати самостійно, що у свою чергу приведе до більш точних результатів. Оволодіння навичками з «конструювання підказок» займає час і вимагає окремого навчання.

Опанування цими навичками ускладнюється тим, що різні програми можуть вимагати по-різному створені запити. Для одних достатньо точно сформульованого запиту створеного природньою мовою, для інших необхідно знання і розуміння окремих команд, які відкривають більш широкий функціонал.

Крім того, для користування програмами на основі ГШІ типу текст-в-зображення необхідним є сучасне аудіо-візуальне обладнання та фінансування щодо закупівлі ліцензій на зазначені програми.

Серед недоліків сучасних програм на основі ГШІ типу текст-в-зображення є те, що багато компаній накладають обмеження на свої програми для запобігання створенню неприйнятних зображень. Це включає в себе фільтри та алгоритми модерації, які блокують створення потенційно неприйнятної контенту. Це зменшує можливості щодо використання подібних програм і у правомірних цілях також.

Ключові слова: ГШІ, текст-в-зображення, заняття, англійська мова, ЗВО.

Сучасний світ переживає час інновацій, і генеративний штучний інтелект (ГШІ) став одним із найважливіших досяг-

нень цього процесу. Його вплив поширюється на різні сфери, від творчих індустрій, де він допомагає створювати, наприклад, нові музичні композиції [1, с. 12] та картини [2, с. 126], до бізнесу, де він оптимізує процеси та допомагає у створенні нових ідей [3, с. 112]. ГШІ має потенціал докорінно змінити наш спосіб роботи, навчання та взаємодії з технологіями.

Взагалі, Генеративний штучний інтелект - це підвид штучного інтелекту, який використовує генеративні моделі для створення тексту, зображень, відео або інших форм даних [4, с. 2]. Тобто, ГШІ навчається на великій кількості даних, таких як тексти, зображення або музика. Він використовує нейронні мережі, аби виокремлювати шаблони і закономірності в цих даних. Після навчання ГШІ може створювати нові, унікальні речі на основі того, чому він навчився, наприклад, писати нові тексти чи створювати нові картини.

Постановка проблеми. ГШІ має також значний вплив і на сферу освіти. Він надає можливість зробити навчання більш персоналізованим, адаптивним та ефективним. ГШІ може допомогти у створенні індивідуальних навчальних планів, навчальних матеріалів та інтерактивних сценаріїв під час вивчення мов. Це технологія, яка може змінити підхід до освіти, зробити її більш ефективною та індивідуалізованою. Проблема ГШІ полягає в тому, що ця технологія є новою і у публічний простір вона потрапила лише декілька років назад. Через це вчені тільки ще починають вивчати та описувати підходи до її застосування в різних сферах життя, в тому числі і в освіті. Тому, для осучаснення методик викладання в ЗВО, включаючи і проведення занять з англійської мови, необхідним є глибокий і всебічний аналіз можливостей і недоліків ГШІ на прикладів існуючих сьогодні програм.

Аналіз останніх досліджень. Дослідження можливостей ГШІ є популярною темою у усьому світі. Вчені, проводять наукові дослідження на кшталт «Від дошки до навчання за допомогою ШІ» стосовно історії введення штучного інтелекту в сферу освіти [5, с. 386]. Вивчаються також і можливості застосування саме ГШІ в сфері освіти. Сандху Р. та інші проаналізували можливий майбутній розвиток застосування інструментів на основі ГШІ у сфері освіти [6]. Проводяться і поглиблені дослідження з більш вузьких питань, наприклад, щодо перетворення вищої освіти у більш персоналізовану за допомогою ГШІ [7] та розробки навчальних планів за допомогою ГШІ [8, с. 227].

Мета статті. Як вже було згадано раніше, можливість застосування ГШІ в сфері освіти потребує всебічного аналізу, який зараз проводиться вченими по всьому світу. Проте, явище ГШІ є занадто обширним і його застосування може змінюватись в залежності від типу ГШІ і предмету, котрий викладається.

Тому мета цієї роботи полягає у висвітленні однієї зі складових цього більш глобального питання, а саме, використання такого типу ГШІ, що перетворює текст у зображення для проведення занять з англійської мови в ЗВО.

Основний матеріал. На сьогоднішній день, генеративний штучний інтелект знаходить широке застосування в програмах, що перетворюють текст у зображення (text-to-image). Ці програми використовують нейромережі для створення зображень на основі текстових описів.

Генеративний штучний інтелект типу текст-в-зображення має потенціал покращити процес проведення занять з англійської мови в ЗВО. Головна можливість його застосування – це самостійне створення викладачем ілюстрацій та візуальних матеріалів. Програми, що перетворюють текст в зображення можуть створювати ілюстрації для нових слів та фраз, з якими повинні ознайомитись студенти. Викладач, зі своєї сторони, матиме можливість заощадити час за рахунок усунення необхідності пошуку ілюстративних матеріалів.

Навчальним матеріалам, котрі генерує ГШІ можна надавати різні рівні складності та деталізації і різну направленість. Це надає можливість викладачу підлаштовувати матеріал заняття під можливості, потреби та інтереси різних груп або окремих студентів.

Проте, ефективне використання ГШІ типу текст-в-зображення на заняттях з англійської мови вимагає від викладачів наявності певних технічних і педагогічних навичок, а також сучасних аудіовізуальних засобів для роботи з програмами.

Однією із головних навичок є *конструювання підказок* (prompt engineering) – це процес проектування, тестування та оптимізації підказок, які надсилаються до штучного інтелекту [9, с. 159]. Відносно викладачів, мета опанування цього навичку полягає в тому, аби вміти формулювати підказки, тобто створювати запити таким чином, аби у відповідь модель штучного інтелекту надавала найбільш точний і бажаний результат.

У приклад програм на основі генеративного штучного інтелекту типу текст-в-зображення можемо навести DALL-E 3, яка вбудована в чат-бот ChatGPT.

Щоб створити зображення в DALL-E 3 в ChatGPT, необхідно надати детальний опис бажаного зображення. Наприклад, при вивченні зі студентами теми овочів одним з можливих запитів є: «Створи зображення тарілки з капустою, баклажаном, броколі, часником, помідором, огірком». Замість вказаних овочів, за необхідності, можна вибрати і будь-які інші. Через деякий час, який займає створення нового зображення, ChatGPT згенерує зображення відповідно до наданого опису. Чим детальнішим буде опис, тим ближче до побажань користувача буде створене ГШІ зображення.

Розуміння того, як конструювати правильні підказки для ГШІ, має прямий вплив на майбутній результат та ефективність і доцільність всього процесу. Правильно створені підказки гарантують, що ГШІ зрозуміє контекст і створить бажане зображення з першого запиту. Погане володіння навичками конструювання підказок призведе до необхідності переробляти запит і створювати нову ілюстрацію по причині недостатньої релевантності першої.

Точно сконструйовані підказки у запитах мінімізують двозначність і допомагають ГШІ генерувати більш точні та релевантні зображення. Тобто, якість вхідних даних безпосередньо впливає на якість вихідних, що робить вміння конструювати

правильні підказки ключовим компонентом у повноцінному використанні генеративних моделей штучного інтелекту.

Проявляється це, наприклад, при персоналізації створюваного ілюстративного матеріалу. Однією з тем під час вивчення нової лексики в будь-якій мові, в тому числі і в англійській, є зовнішність людини. Якщо викладач проводить заняття з англійської мови для студентів, яким подобається японська мультиплікація, а у підручнику є лише зображення в фото-реалістичному стилі, то, відповідно, при створенні запиту для ГШІ можна задати і необхідний стиль. Один із варіантів такого запиту: «Створи деталізоване зображення людського обличчя. Стиль японської анімації». При необхідності, створений запит можливо уточнити задля більш конкретного результату. Наприклад, при бажанні відтворити японську анімацію певних часів, друге речення в запиті можна змінити додавши підказку: «Стиль японської анімації 90-х років» (Рис. 1). Після такого уточнення, нове зображення буде створене з урахуванням більшої кількості інформації, і буде точніше відповідати побажанню користувача.

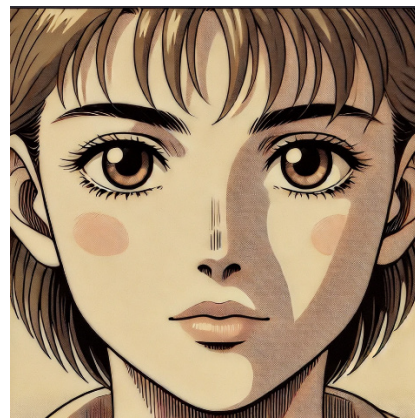


Рис. 1. Зображення отримане під час скорегованого запиту

Відсутність відповідних навичок у викладача може призвести до неефективного використання ГШІ. Як приклад можемо зазначити, що остаточний запит, на основі якого DALL-E 3 створив зображення на рис. 1 виглядає наступним чином: «Створи деталізоване зображення людського обличчя. Стиль японської анімації 90-х років». В цьому запиті не вистачає підказок для ГШІ стосовно різних деталей. Наприклад, статі людини, яка повинна бути на ілюстрації. DALL-E 3 самостійно вирішив створити ілюстрацію з людиною жіночої статі. Якщо б серед необхідної для вивчення студентами лексики був волоссяний покрив на обличчі у чоловіків, то така ілюстрація унеможливила б його демонстрацію, що призвело б до виникнення необхідності створення іншої ілюстрації.

Таким чином, викладач, який усвідомлює важливість правильного конструювання підказок для ГШІ і володіє відповідними навичками може ефективно і швидко створювати та персоналізувати ілюстративні матеріали для занять. Низький рівень володіння такими навичками може зробити процес створення ілюстрацій більш довгим та неефективним, а результат нерелевантним.

Іншою складністю відносно конструювання підказок є те, що різні програми на основі ГШІ типу текст-в-зображення вимагають різних підходів.

Так, для роботи з DALL-E 3 достатньо точно і правильно сконструйованих запитів створених природною мовою, на кшталт вказаних раніше. Але, наприклад, для задіяння просунутих функцій в іншій програмі такого ж типу Midjourney, необхідно вивчити та опанувати спеціальні команди, на кшталт «--no plants», яка усуне з зображення всі рослини (замість plants можна поставити будь-який інший тип об'єктів), «--style <4a, 4b, or 4c>» аби переключатись між різними підверсіями програми четвертої версії та інші [10].

Існування подібних розбіжностей в області користування програмами на основі ГШІ типу текст-в-зображення вимагає належного навчання викладачів з ціллю опанування всіх можливостей вибраної для роботи програми.

Крім того, для застосування ГШІ типу текст-в-зображення необхідною є наявність у викладача сучасних аудіовізуальних засобів, які можуть відтворити зображення. Не кожний заклад вищої освіти має належне фінансування і може надати викладачу необхідні аудіовізуальні засоби та провести тренінги з їх використання.

Питання участі університету також виникає на етапі надання ліцензій на використання сучасних програм на основі ГШІ типу текст-в-зображення. Згаданий раніше DALL-E 3 в ChatGPT можна використовувати безкоштовно, хоча і з деякими обмеженнями. Проте, багато програм, наприклад Midjourney, вимагають придбання ліцензії з щомісячною оплатою, що для багатьох викладачів унеможливить їх використання.

Серед недоліків ГШІ типу текст-в-зображення можемо зазначити те, що багато компаній обмежують можливості своїх програм на основі ГШІ з ціллю запобігти створенню неприйнятних зображень. Ці обмеження полягають в наявності фільтрів та алгоритмів модерації контенту, котрі існують задля блокування або позначення контенту, який може бути неприйнятним. Наприклад, моделі штучного інтелекту навчаються розпізнавати і запобігати створенню відвертих, насильницьких або інших неприйнятних зображень.

Однак такі обмеження можуть ненавмисно перешкоджати правомірному використанню цих програм. Створюючи зображення для уроків англійської мови, викладачі в ЗВО можуть зіткнутися з тим, що їхні можливості будуть обмеженими. Фільтри ГШІ можуть помилково відхилити правомірний запит на створення певної ілюстрації. Наприклад, викладач, який хоче створити зображення певних історичних або літературних сцен, може зіткнутися з труднощами, якщо вони будуть помилково визначені як неприйнятні.

Як приклад описаного вище можемо привести інший запит для DALL-E 3. Звертаючись до вже згаданої теми людської зовнішності, DALL-E 3 створив ілюстрацію (рис. 1) у відповідь на вказаний раніше запит «Створи деталізоване зображення людського обличчя. Стилль японської анімації 90-х років». Проте, людське тіло не обмежується лише обличчям, що відноситься і до лексики, яку необхідно вивчати студентам. На інший запит «Створи анатомічно правильне зображення чоловічого торсу. Реалістичний стилль.» DALL-E 3 відповів наступним повідомленням про помилку: «Я не зміг згенерувати анатомічне зображення чоловічого торсу через проблеми із запитом. Якщо ви бажаєте, ми можемо спробувати інший підхід або змінити деталі. Дайте мені знати, як ви хотіли б діяти далі!». Проте, у випадку уточнюючого запиту направленного на уникнення можливої демонстрації оголеного чоловічого торсу «Створи

зображення чоловіка в діловому костюмі, на зображенні повинно бути видно лише торс. Реалістичний стилль.» зображення було створено успішно (рис. 2).



Рис. 2. Зображення отримане після виправленого запиту

З цього випливає, що обмеження щодо створення зображення у відповідь на початковий запит не є технічними, а існують через модераційні запобіжники. Крім того, на сайті програми прямо зазначено про ряд обмежень, щодо генерування зображень [11].

Хоча такі запобіжники мають важливе значення для дотримання етичних стандартів, вони можуть обмежувати ілюстрування деяких тем і можливості викладачів. Баланс між забезпеченням безпеки та наданням можливостей є постійним викликом для розробників ГШІ. Компаніям необхідно вдосконалювати свої алгоритми, щоб зменшити кількість помилкових спрацьовувань у відповідь на правомірні запити, зберігаючи при цьому надійний захист від створення дійсно шкідливого контенту.

Висновки. Використання програм на основі генеративного штучного інтелекту типу текст-в-зображення має значні перспективи щодо покращення проведення занять з англійської мови в ЗВО. Проте від викладачів, які будуть користуватись ГШІ, вимагається володіння навичками правильного конструювання підказок для правильної роботи з програмами. Без вірно створених запитів ефективність користування ГШІ значно знизиться, а результати погіршаться. Ситуацію ускладнює і те, що в різних програмах підказки можуть конструюватись по-різному і в деяких випадках необхідним є володіння особливими командами.

Іншою важливою проблемою є надання університетами необхідних аудіо-візуальних засобів, котрі нададуть можливість працювати з програмами на основі ГШІ типу текст-в-зображення, та придбання ліцензій відповідних програм.

Серед недоліків у застосуванні можемо відзначити наявність фільтрів та алгоритмів, котрі унеможливають створення ілюстрацій з деяких тем.

Перспективи майбутніх досліджень вбачаються в вивченні можливих способів подолання окреслених проблем та імплементації ГШІ типу текст-в-зображення в ЗВО України та інших країн.

Література:

1. Atanacković D. Artificial Intelligence: Duality in Applications of Generative AI and Assistive AI in Music. *INSAM Journal of Contemporary Music, Art and Technology*. 2024, pp. 12–31.

2. Rani S. et al. The Role of Artificial Intelligence in Art: A Comprehensive Review of a Generative Adversarial Network Portrait Painting. 2023, pp. 126–135.
3. Іваненко А., Пічик К. «Генеративні моделі штучного інтелекту як ефективний інструмент для оптимізації бізнес-процесів». *Емпірію*. 2024. Т.1, № 1, С. 112–121.
4. Pinaya W. et al. "Generative AI for Medical Imaging: extending the MONAI Framework". arXiv preprint, 2023, pp. 1–23.
5. Güner H. et al. From Chalkboards to AI-Powered Learning: Students' Attitudes and Perspectives on Use of ChatGPT in Educational Settings. *Educational Technology & Society*, vol. 27, no. 2, 2024, pp. 386–404.
6. Sandhu R. et al. An Introduction to Generative AI Tools for Education 2030. *Integrating Generative AI in Education to Achieve Sustainable Development Goals*. 2024, pp. 1–28.
7. Blessing M. Leveraging Generative AI for Personalized Learning in Tertiary Education. 2024
8. Tariq M. Generative AI in Curriculum Development in Higher Education. *Impacts of Generative AI on Creativity in Higher Education*. 2024, pp. 227–258.
9. Lemeš S. Prompt Engineering. Artificial intelligence in industry 4.0: The future that comes true. 2024, pp. 159–170.
10. Midjourney Parameter List. *Midjourney Documentation*. URL: <https://docs.midjourney.com/docs/parameter-list> (дата звернення: 26.10.2024).
11. DALL-E 3. *OpenAI*. URL: <https://openai.com/index/dall-e-3/> (дата звернення: 26.10.2024).

Domarenko M. The use of software with text-to-image generative AI for teaching English in higher education institutions

Summary. The article provides a brief explanation of generative AI in general and text-to-picture generative AI in particular, as well as reviews the latest research on its application in education.

As a new and complex phenomenon, the utilization of generative AI can vary depending on the field and type of generative AI itself. For this reason, the article focuses on a specific part of this global issue, namely the analysis of the use of text-to-image software for English language teaching in higher education institutions.

As generative AI has only recently become widespread, its use in education is only just beginning to take place.

The main way to utilize text-to-image software for English language classes in higher education institutions is to create illustrations and various visual materials, which can help learn new vocabulary and visualise complex concepts and ideas.

In order to use these programmes effectively, teachers need to master the skills of prompt engineering that will help them get relevant results.

To perform its task, a generative AI needs precise instructions. The more precise and complete the prompts are, the less the generative AI will have to guess, which in turn will lead to more accurate results. Mastering the skills of «prompt engineering» takes time and requires separate training.

Mastering this skill is complicated by the fact that different software may require differently constructed prompts. For some, a well-formulated prompt created in natural language is enough, while others require knowledge and understanding of certain commands that open up wider functionality.

In addition, modern audio-visual equipment and funding for the purchase of licences for these products are required to use text-to-image software.

One of the disadvantages of modern text-to-image software is that many companies impose restrictions on their products to prevent the creation of unacceptable images. This includes filters and moderation algorithms that block the creation of potentially offensive content. This is also a limitation of the software's capabilities in the context of legitimate use.

Key words: generative AI, text-to-image, teaching, English language, HEI.