
ВИКОРИСТАННЯ МОЖЛИВОСТЕЙ НЕЙРОМЕРЕЖ ТА ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ СТВОРЕННЯ МЕДІАКОНТЕНТУ



Наталія Чернишенко

ПВНЗ «Харківський міжнародний медичний
університет»
Бібліотека
ORCID iD 0000-0002-8750-6527

Розглянуто можливості різних нейронних мереж як інструментів для створення графічного, текстового та аудіовізуального контенту; наведено приклади поєднання можливостей різних нейромереж та штучного інтелекту в межах одного проекту бібліотеки.

Ключові слова: нейромережі, штучний інтелект (ШІ), ChatGPT, графічний контент, аудіовізуальний контент, медіаконтент, штучний інтелект.

Розвиток технології штучного інтелекту (ШІ) продовжує набирати обертів і охоплює різні сфери людської діяльності. Вона набула шаленої популярності в усьому світі, не залишила осторонь і нашу країну, особливо після виходу ChatGPT на український ринок. Ця технологія базується на використанні нейронних мереж, які моделюють функціонування людського мозку. Такі мережі складаються з великої кількості штучних нейронів, які взаємодіють між собою і призначені для обробки вхідних даних [1].

На сьогодні існує широкий спектр нейромереж, розроблених на основі штучного інтелекту, які здатні виконувати різноманітні завдання. Серед них важливе місце займають можливості розпізнавання мовлення, генерування перекладу та переформулювання текстової інформації, а також пошуку та відбору даних. Технологія штучного інтелекту також використовується для створення музики та зображень, включаючи портрети, що генеруються на основі заданих текстових описів [4]. Безумовно, такі можливості нейромереж та штучного інтелекту є просто незамінними при створенні високоякісного бібліотечного медіаконтенту, який зможе привернути увагу користувача та допоможе бібліотекам вийти на новий рівень комунікації.

У цій статті ми дослідимо ресурси, які базуються на штучному інтелекті та спрямовані на створення та редагування графічних зображень, розглянемо використання чату GPT для ідеї генерації та написання сценаріїв, а також створення аватарів з аудіо компонентом. Ми також продемонструємо на прикладах, як можна комбінувати можливості різних нейромереж у бібліотечних проєктах.

Наразі для швидкого та ефективного генерування високоякісних зображень для будь-яких бібліотечних цілей найбільш придатними та універсальними серед багатьох є наступні нейромережі: Leonardo AI, Midjourney, DALL-E, DreamStudio, Adobe Firefly. Створення візуального контенту в цих нейромережах відбувається за текстовим запитом («prompt»), який пише людина. Тому важливо підготувати деталізований і розгорнутий текстовий запит англійською мовою, що буде впливати на якість отриманого візуального результату. Щоб спростити процес написання промту можна скористатися послугами ще однієї нейромережі Adobe Firefly prompt generator.

Leonardo AI (<https://app.leonardo.ai/>) – це інноваційний ресурс, який відкриває широкі можливості для розробки різних застосувань штучного інтелекту. Ця платформа створена та підтримується компанією OpenAI, визнаною лідером у сфері інтелектуальних систем. На Leonardo AI користувач може вибрати модель, налаштування та створити запит, будь-то на основі ідеї чи зображення. Платформа працює не тільки на власних моделях, але й на спеціалізованих моделях, які розробляє спільнота дизайнерів. На день Leonardo AI надає 150 токенів, 4 токена коштує генерування за одним запитом (4 зображення). Завдяки Leonardo AI, бібліотеки можуть створювати візуальний контент для своїх проєктів (аватари, банери, рекламний контент для соцмереж, постери та матеріали для реклами книг та заходів), без необхідності глибоких знань графічного дизайну та суттєво економити час для його створення [5].

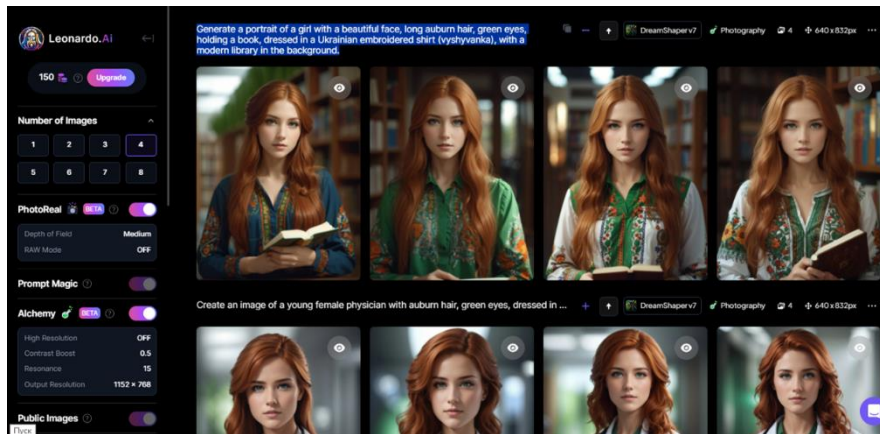


Рис. 1, 2. Приклади згенерованих зображень за допомогою Leonardo AI

DALL-E (<https://labs.openai.com/>) – нейронна мережа, від компанії OpenAI, розробників ChatGPT, яка дозволяє користувачам створювати зображення на основі текстового запиту. За один запит може бути створено чотири зображення, з різними темами та стилями – від реалістичних до абстрактних. Якість результатів безпосередньо залежить від якості текстового запиту, більш точний запит збільшує ймовірність того, що DALL-E створить потрібні зображення [4].

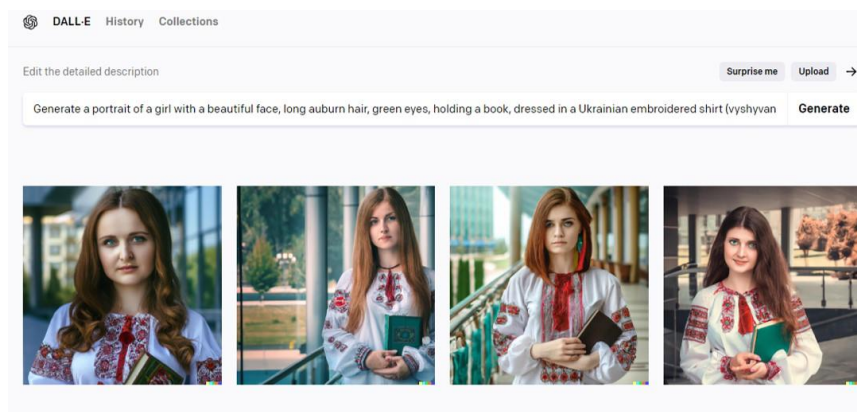


Рис. 3. Приклад згенерованого зображення за допомогою DALL-E

DreamStudio (<https://beta.dreamstudio.ai/generate>) – інструмент для генерування зображень на базі штучного інтелекту, який дозволяє користувачам створювати високоякісні та неповторні графічні елементи без використання графічних редакторів. Сервіс є платним, але при реєстрації можна отримати двадцять п'ять кредитів, які дозволяють створити приблизно тридцять фотографій.

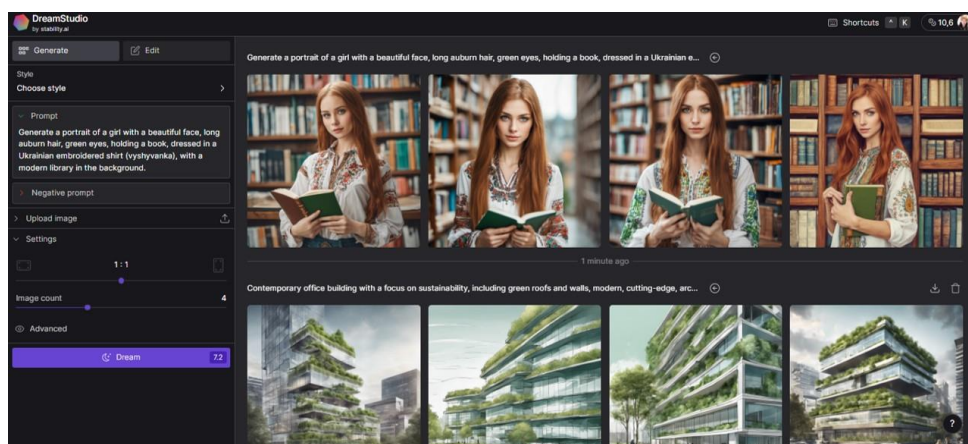


Рис. 4. Приклад згенерованого зображення за допомогою DreamStudio

Adobe Firefly (<https://firefly.adobe.com/>) – це генеративний інструмент штучного інтелекту для створення унікальних та оригінальних дизайнів, який дозволяє користувачам об'єднувати різні елементи в одному кадрі. Користувачі можуть налаштовувати кольорову палітру, змінювати форму і розмір елементів, додавати або видаляти елементи та текстури, використовуючи текстові вказівки.

Після створення зображення його можна додатково обробити за допомогою вбудованих інструментів, не потрібно вводити додаткові текстові запити. Генерація зображень з тексту є популярною, оскільки вона проста. Чим більш точний опис, тим кращий результат. Adobe Firefly розуміє українську мову, що дозволяє користувачам вводити текстові запити без використання перекладача. Ця модель Firefly також надає можливість створювати стилі і текстури для тексту.

Наразі Adobe Firefly доступний та дозволяє 25 безкоштовних щомісячних генеративних кредитів. Важливо відзначити, що модель розпізнає український прапор і, в більшості випадків, генерує його правильно. Це робить цей ресурс ідеальним для створення патріотичних зображень [2].

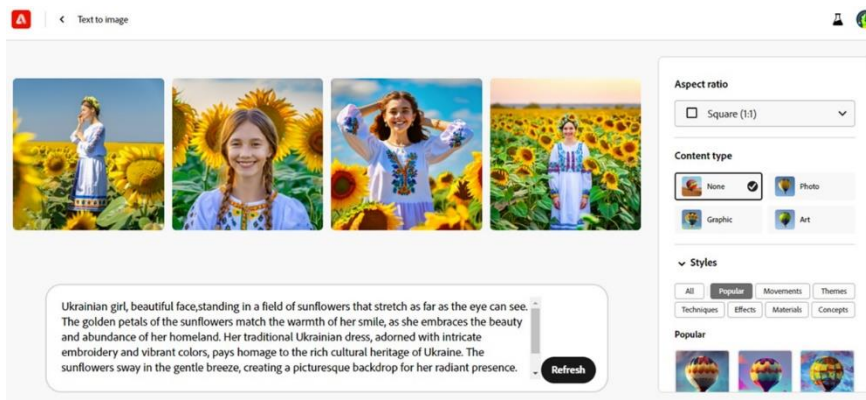


Рис. 5. Приклад згенерованого зображення за допомогою Adobe Firefly

Існують мовні моделі штучного інтелекту, такі як ChatGPT, Character AI, BLOOM, які здатні створювати сценарії, вірші, генерувати ідеї, що можуть використовуватися для початку творчого процесу.

ChatGPT (<https://chat.openai.com/>) – це штучний інтелект, заснований на моделі GPT, яка є неймерережевою архітектурою на основі трансформера. Модель GPT була спеціально навчена на великому обсязі текстових даних для розуміння мови та генерації відповідей на запитання.

За даними аналізу, проведеного швейцарським банком UBS, встановлено, що ChatGPT стрімко розвивається. За цими даними, у січні, лише через два місяці після свого запуску, ChatGPT досяг 100 мільйонів активних користувачів. Для порівняння, TikTok знадобилося дев'ять місяців для досягнення того ж показника, тобто 100 мільйонів активних користувачів [3].

Штучний інтелект, зокрема ChatGPT, дозволяє генерувати ідеї, писати сценарії, вірші для віртуального контенту з урахуванням текстових запитів та аналізу текстової інформації.

Studio.d-id (<https://studio.d-id.com/>) є високорівневою технологією, розробленою для створення аватарів, які включають аудіоелемент. Це означає, що аватар може не лише візуально відтворювати образ особи, але й генерувати мовлення, відповідне цьому образу. Ця можливість робить її корисною для створення інтерактивних розмовних інтерфейсів, віртуальних помічників та інших застосувань, де важливо забезпечити аутентичний досвід спілкування.

Неймерережа Studio.d-id створює аватари, які можуть говорити різними мовами, що є надзвичайно корисним інструментом для бібліотек ЗВО з іноземним контингентом.

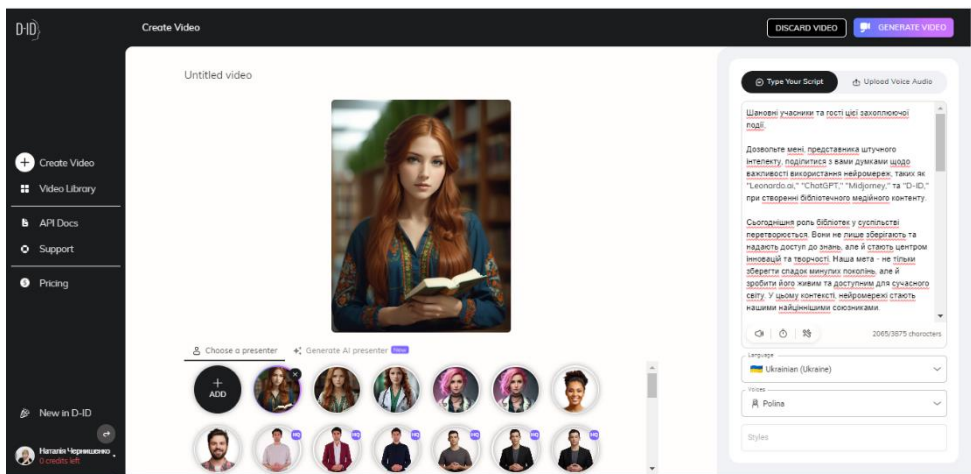


Рис. 6. Приклад згенерованого зображення за допомогою Studio.d-id

Ефективним у роботі бібліотеки є поєднання та доповнення можливостей різних нейромереж та штучного інтелекту при створенні одного проєкту. Наприклад, цього року у Харківському міжнародному медичному університеті відбулася студентська науково-практична конференція «Я обираю здоров'я». Бібліотека підготувала до конференції список літератури: «Як війна впливає на здоров'я українців», який був представлений із застосуванням сучасних нейронних мереж та чату GPT. <https://drive.google.com/drive/my-drive?q=type:video%20parent:0ABtuOrh0AwEDUk9PVA>. При цьому було використано нейромережу Midjourney, який створив портрет аватару, ChatGPT згенерував текст звернення аватару двома мовами (українською та англійською), а нейромережа Studio.d-id згенерувала аудіо-візуальний аватар, який і презентував список літератури під час конференції.

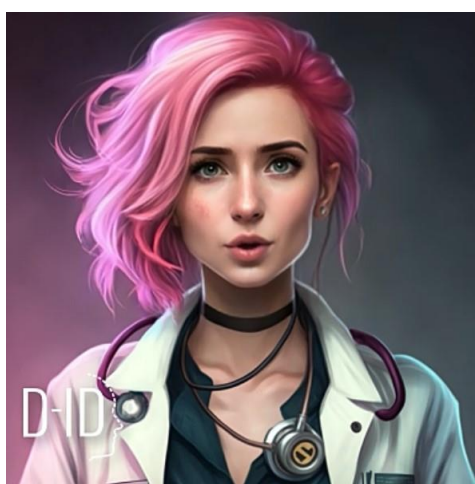


Рис. 7. Приклад згенерованого аватару за допомогою Studio.d-id

Ще приклад, до 209-ї річниці від дня народження Тараса Шевченка було створено цілу низку відео контенту, де штучний інтелект читає вірші Кобзаря. У цьому випадку аватар був згенерований нейромережою Studio.d-id, а аудіозапис створений власноруч <https://youtu.be/kA1V9vejyLA>.



Рис. 8. Приклад згенерованого автатару за допомогою Studio.d-id

Перевагою застосування нейронних мереж та штучного інтелекту є:

- значне прискорення процесу генерування візуального контенту і можливості комбінувати різні програми та доповнювати й вдосконалювати результати кожної з них;
- створення візуального контенту за текстовим запитом, тобто користувачу не потрібно навчатися малювати, що потребує суттєвої витрати енергії, часу та ресурсу;
- урізноманітнює види подачі інформації, робить контент бібліотеки цікавим і сучасним;
- надає можливість створення аудіовізуального контенту будь-якою мовою.

Недоліки застосування нейронних мереж та штучного інтелекту:

- етичні виклики, пов'язані з штучним інтелектом, включають створення фальшивих зображень і діпфейків, що вже призвело до скандалів, таких як зображення Папи Римського Франциска, який одягнутий у пуховик Balenciaga та арешт Дональда Трампа, де використання ШІ призвело до надмірно реалістичних образів;

- авторське право стоїть перед викликом, оскільки більшість систем штучного інтелекту, включаючи Midjourney, DALL-E та інші, базуються на авторському контенті в базах даних. Отримані зображення є колажем робіт різних авторів, які не надавали згоду на використання своїх творінь для навчання штучного інтелекту [2].

Підсумовуючи сказане, варто зробити висновок, що технології ШІ та нейромережі дійсно мають великий потенціал. Вони надають сучасним бібліотекам безліч можливостей створювати різноманітний та високоякісний віртуальний контент, оптимізований для своєї цільової аудиторії. Потрібно тільки бажання та час, щоб опанувати цілу низку ресурсів.

Література:

1. Божко Т., Ареф'єв В. Нейронні мережі як інструмент графічного дизайну. *Вісник КНУКІМ. Серія «Мистецтвознавство»*. 2023. № 48. С. 125–135. DOI: [10.31866/2410-1176.48.2023.282475](https://doi.org/10.31866/2410-1176.48.2023.282475).

2. Дьякова А. Школа бібліотечного журналіста. AI+Art: нові горизонти у творчості бібліотекаря. *YouTube*. 2023. URL: <https://www.youtube.com/watch?v=rqafSycYgSE> (дата звернення: 19.09.2023).

3. Чат-бот ChatGPT: можливості й загрози. *Медіамейкер*. URL: <https://mediamaker.me/chat-bot-chatgpt-shho-potribno-znaty-vydavczyam-569/> (дата звернення: 20.09.2023).

4. Ямковенко В., Тітова Л. Сервіси для роботи з графічними зображеннями на базі штучного інтелекту. *Наука. Освіта. Молодь* : зб. матеріалів XVI Всеукр. наук. конф. студентів та молодих науковців (Умань, 11 травня 2023 р.) / МОН України, Уманський держ. пед. ун-т ім. Павла Тичини ; за ред. О. І. Безлюдного. Умань : Візаві, 2023. С. 393–395.

5. Leonardo AI. Безкоштовна нейромережа для генерації зображення. *Знайшов*. URL: <https://tinyurl.com/2sw9zsjc> (дата звернення: 24.09.2023).

6. Midjourney – AI для генерації зображень. Як з ним працювати. *AIN.UA*. URL: <https://ain.ua/2023/02/24/shho-take-midjourney-i-yak-z-neyu-praczuvaty/> (дата звернення: 17.08.2023).

USING THE OPPORTUNITIES OF NEURAL NETWORKS AND ARTIFICIAL INTELLIGENCE TO CREATE MEDIA CONTENT

Nataliia Chernyshenko

Private Institution of Higher Education «Kharkiv International Medical University»

Library

ORCID iD 0000-0002-8750-6527

The possibilities of various neural networks as tools for creating graphic, textual, and audiovisual content are examined. Examples of combining the capabilities of different neural networks and artificial intelligence within a single library project are provided.

Key words: neural networks, artificial intelligence (AI), ChatGPT, graphic content, audiovisual content, media content, digital technologies.