

ШТУЧНИЙ ІНТЕЛЕКТ: ОСОБЛИВОСТІ ВИКОРИСТАННЯ В НАУКОВИХ ДОСЛІДЖЕННЯХ

Яценко Оксана

асистент

oksana@zu.edu.ua

Мінгальова Юлія

асистент

mailto:mingalyovay@gmail.com

Яценко Олександр

асистент

sas@zu.edu.ua

Кафедра комп'ютерних наук та інформаційних технологій

Торгонська Анастасія

фахівець з обробки електронної інформації

Відділ обслуговування комп'ютерних мереж

anastasia.torgonskaja@gmail.com

Житомирський державний університет імені Івана Франка

Історія розвитку штучного інтелекту(ШІ) триває вже більше 70 років. За цей час технологія встигла пройти декілька кризових періодів, що характеризувались зменшенням фінансування та інтересу до досліджень, та були названі зима ШІ, і вийти на пік популярності, масової залученості після випуску чат-бота OpenAI AI Generative Pre-Trained Transformer 3 (GPT-3) в листопаді 2022 року. Демонструючи вражаючі результати в генерації текстів і зображень, ChatGPT, що зараз має понад 180 мільйонів користувачів щомісяця та 100 мільйонів активних користувачів щотижня [1], також стимулював нову хвилю дискусій і обговорень пов'язаних з етичністю, авторським правом, плагіатом.

Світова і українська наукова спільнота періодично зіштовхуються з різноманітними проявами академічної недоброчесності. Так німецький політик К. Гуттенберг в 2011 році був змушений подати у відставку після викриття плагіату в його дисертації. Схожі звинувачення щодо плагіату і відсутності цитувань першоджерел, висувались на адресу першої леді Кореї Кім Кен Хі. Академічна недоброчесність в «побутових» її проявах найчастіше залишається безкарною, проте у випадку відомих осіб нерідко стає інструментом для

маніпуляцій, нанесення репутаційних втрат, за якою часто стоїть виразна політична вмотивованість [2].

Постає питання: варто розглядати ChatGPT в контексті допоміжних інструментів для полегшення написання наукових робіт, чи як новий фактор що призведе до урізноманітнення проявів академічної недоброчесності?

Під час створення наукового або творчого контенту людина опрацьовує власний досвід, компілюючи результат на основі попередніх знань та сприйняття, тобто буквально те що створено іншими в тому чи іншому масштабі здійснює вплив на нову ідею. Модель ChatGPT працює за подібними алгоритмами. Використовуючи великі масиви тренувальних даних за спеціальними рівняннями з безліччю параметрів [3], модель намагається згенерувати валідну відповідь на поставлене запитання, яка буде заснована на вхідній інформації, на якій навчалась модель. Відмінність полягає в тому, що ChatGPT підсумовує і узагальнює вже відомі ідеї, статистично вираховуючи кожне наступне слово тексту, в той час як людина намагається прийти до нових висновків і знань.

Ще однією проблемою з якою може зіткнутися автор, при написанні наукової роботи за допомогою штучного інтелекту, неактуальність даних. Справа в тому, що масиви інформації включають лише конкретний часовий період, не підтримуючи зв'язку з новими дослідженнями. В наш час, коли кожного дня відбуваються оновлення інформації і модернізація технологій, це може стати великою проблемою в роботах, що написані засобами штучного інтелекту. GPT-4, як правило, не знає про події, які відбулися після того, як завершилась передача переважної більшості його даних (вересень 2021 року), і не вчиться в режимі реального часу.

Генеруючи текст, модель ШІ використовує публічні дані, серед яких і наукові праці певних авторів. Він не копіює та не вставляє інформацію безпосередньо з джерела, але може скопіювати ту саму ідею та перефразувати текст. Вміст може бути схожим на деяке існуюче джерело та реструктуровану його версію. Тому переформулювання речень для зменшення відсотка плагіату не можна вважати прийнятним у наукових дослідженнях. Якщо ми визначаємо «плагіат» як звичайне копіювання чиеїсь роботи, лише перефразування написаного, незалежно від використаного методу, і без додавання особистого, є порушенням академічної доброчесності. З цієї причини редактори журналів повинні використовувати програми для виявлення фрагментів наукової роботи, що згенеровані за допомогою ШІ, щоб краще виявляти плагіат [4].

Але навіть уникнувши плагіату, використання ChatGPT в наукових дослідженнях, призводить до неочевидного наслідку: опрацьовуючи інформацію самостійно, автор повинен додавати цитування на конкретні джерела, ChatGPT розмиває межі різних робіт що фактично не вимагає цитування і посилання на них. Як результат певний відсоток цитувань буде втрачено, а рейтинг перегляду статей та їх індекс цитування, закономірно, стане меншим [4].

Інша проблема висновків, які генерує ChatGPT керуючись доступними даними, – це їх неправдивість. «Брехати» його змушує явище назване

галюцинацією ШІ. Галюцинації виникають з різних причин, серед яких: брак розуміння реального світу, помилки при розробці та написанні програмного забезпечення, некоректне або неоднозначне запитання, специфічні напрями науки, або технологій [5]. Основна задача ChatGPT надати відповідь користувачу, крім випадків коли запит суперечить політиці використання, що не гарантує при цьому достовірність інформації. Описана проблема не робить ChatGPT непорідним, але вказує на його ненадійність.

ChatGPT часто уникає відповіді на питання що стосуються конфліктних чи спірних ситуацій, або певних осіб (популярних зірок, відомих діячів, тощо). Крім того технологія може бути умовно упередженою діючи на користь певних компаній монополістів, що нею володіють.

Ситуація прозорого використання ChatGPT чи іншої нейронної мережі в наукових роботах виявляється також доволі спірною. Штучний інтелект не може брати на себе відповідальність за коректність висловлювань, їх достовірність чи унікальність. Технологія все ще залишається хоч і потужним, але керованим людиною інструментом, що створює вплив суб'єктивного бачення дійсності, та перспективи підтасовування фактів і висновків до необхідних автору. Якими потужними не здавалися б можливості будь якого штучного інтелекту, він все ще залишається суб'єктивним, позбавленим особистості та нездатним відстоювати певну точку зору.

З метою обмеження використання ШІ в наукових роботах існують сервіси, подібні до сервісів перевірки на плагіат, що досліджують текст на використання ШІ. При цьому якщо плагіат можна відслідкувати за першоджерелами, то присутність ШІ вираховується за декількома параметрами, серед яких алгоритм схожий до принципу створення тексту оригінальною нейронною мережею, тобто програма вираховує наскільки статистично вірогідно використання конкретного наступного слова, чим більше співпадінь, тим вища вірогідність авторства ШІ, оскільки для людини непритаманна риса мислити такими категоріями.

Відповідаючи на критику зі сторони вчительської спільноти Генеральний директор OpenAI Сем Альтман сказав, що OpenAI експериментуватиме з технологіями водяних знаків та іншими техніками для позначення контенту, створеного ChatGPT, але він застерігає до уникнення залежності від цих інструментів, оскільки зробити їх ідеально точними навряд чи вийде. «Люди визначають, яку частину тексту їм потрібно змінити. Будуть інші речі, які змінять виведений текст» [6]. Будучи позбавленими 100%-ї впевненості в використанні генеративних матеріалів наукові видання навряд чи можуть відхиляти публікації, керуючись лише певним ризиком участі ШІ.

Незважаючи на описані недоліки нейронні мережі, зокрема ChatGPT знаходяться лише на початку свого розвитку, поступово еволюціонуючи, вони інтегруються до існуючих бізнес, інтернет і наукових середовищ та набувають для них все більшого значення. Вплив ШІ, негативні та позитивні наслідки можна буде оцінити лише з часом. Нові технології на всіх етапах розвитку людства часто лякали і викликали відторгнення у широкого загалу. Проте це не

зупиняє прогрес і врешті решт технологія пройшовши кризу входить до повсякденного життя людини. Отже заборони і страх, не зможуть спинити використання технологій ШІ в різноманітних сферах, включаючи наукову, тому важливо знайти шляхи врегулювання використання штучного інтелекту в дослідницькій та науковій роботі.

Список використаних джерел

2. Shubham Singh. ChatGPT Statistics (June 2024) – Users, Growth & Revenue. Demandsage. URL: <https://www.demandsage.com/chatgpt-statistics/#:~:text=ChatGPT%20has%20over%20180%20million,500%20companies%20are%20using%20ChatGPT> (date of access: 20.05.2024).
3. Максим Мацала. Академічна (не)доброчесність: що з нею не так? URL: <http://surl.li/ueaqo/> (дата звернення: 20.05.2024).
4. How should AI systems behave, and who should decide? URL: <https://openai.com/blog/how-should-ai-systems-behave> (date of access: 20.05.2024).
5. Salvagno, M., Taccone, F. S & Gerli, A. G. Can artificial intelligence help for scientific writing? Crit Care. 2023. Vol. 27, no. 75. URL: <https://doi.org/10.1186/s13054-023-04380-2> (date of access: 20.05.2024).
6. Katie Rees. Is ChatGPT Trustworthy, or Does ChatGPT Lie? MakeUseOf. URL: <https://www.makeuseof.com/does-chatgpt-lie/> (date of access: 20.05.2024).
7. Aaron Mok. CEO of ChatGPT maker responds to schools' plagiarism concerns: «We adapted to calculators and changed what we tested in math class». Business Insider. URL: <https://www.businessinsider.com/openai-chatgpt-ceo-sam-altman-responds-school-plagiarism-concerns-bans-2023-1?international=true&r=US&IR=T> (date of access: 20.05.2024).