

Вікторія ПОЛУЦИГАНОВА

доктор філософії з кібербезпеки та захисту інформації,
асистент кафедри інформаційної безпеки

Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського» (Київ)
medvika@ukr.net

ПРОБЛЕМИ ДЕТЕКЦІЇ ВИКОРИСТАННЯ ГЕНЕРАТИВНОГО ШІ В КВАЛІФІКАЦІЙНИХ РОБОТАХ

На сьогоднішній день генеративний штучний інтелект все частіше використовується в освітньому процесі. Для викладачів – це спосіб краще оброблювати та адаптувати матеріали, які можна використовувати для підвищення власної кваліфікації та впровадження нових знань для студентів. У свою чергу для студентів – це спосіб швидкого пошуку та обробки інформації, інструмент візуалізації тощо. Але, нажаль, існує і негативні наслідки постійного використання штучного інтелекту в якості асистента в пошуку знань та вирішенні задач. Деякі з них полягають у зниженні критичного мислення, довіра до наведених фактів, зниження власної розумової діяльності, особливо, коли це стосується написання текстів та виконання завдань. Тобто, ШІ як нова технологія зробила революцію в більшості сфер життя людей, але через не регламентованість використання часто застосовується для зловживання та махінацій з боку здобувачів освіти [1, 2].

Для попередження надмірного використання інструментів ШІ, особливо при написанні різного роду кваліфікаційних робіт, необхідно розробити шляхи для регламентування цього застосування. Одним з них є детекція використання ШІ. Але потрібно розуміння, що якщо як саме це робити, адже на відміну від плагіату, генеративні ШІ генерують достатньо високої якості тексти, що майже завжди проходять перевірку на плагіат, а точніше, перевірку на співпадіння тексту [2, 3, 4]. Потрібно розроблювати правила, які за якими можна було б відслідковувати таке зловживання. Але з іншої сторони, якщо ШІ допомагає генерувати тексти які мають формальні ознаки і більше втомлюють людину, через те що не мають змістове навантаження, наприклад, встановлені конструкції та фрази, що характерні для кваліфікаційних робіт, то в даному випадку не має сенсу намагатися визначити застосування ШІ. Проблеми виникають, коли здобувачі використовують дані інструменти для фальсифікації та фабрикації даних для своїх досліджень, що в подальшому призводить до погіршення якості самих кваліфікаційних робіт, а головне здобувачі

отримують диплом за роботу, яку вони по факту не провели. Це призводить до проблеми науки заради науки і нових досягнень у розвиток суспільства не несуть.

На зараз надійних способів детекції використання ШІ у дослідженнях не розроблено, хоча вже є деякі гіпотези, щодо можливості виявлення основних ознак. По-перше, які б не були здібні здобувачі, все одно людський фактор накладає такі відбитки як мовленнєві та граматичні помилки, чого в більшості випадків не має в текстах згенерованих за допомогою ШІ. По-друге, стиль написання також може видавати використання ШІ, адже від доволі характерний для кожної LLM (Large Language Model).

Шляхами вирішення даних проблем може бути розробка надійного інструменту для детекції використання ШІ в кваліфікаційних роботах, що є достатньо складною задачею. До того, як і з визначенням тестового співпадіння не означає плагіат за правильного посилання, використання згенерованого тексту не є проблемою у використанні в текстах. Проблема виникає, якщо результати роботи ШІ виставляють як власні здобутки, а особливо непідтверджені згенеровані дані. За умови розробки надійного інструменту детекції, підхід, що призводить до заборони використання ШІ в текстах є практичним і не завжди доцільним. Інший підхід є більш реальним в рамках сучасних реалій. Це проводити курси з використання ШІ та долучати освітні компоненти, для того, щоб здобувачі розуміли як якісно використовувати наявні інструменти ГШІ [5, 6]. Тобто по суті навчити їх використовувати та робити постобробку того що видає ШІ. Це дасть змогу не обтяжувати здобувачів при написанні кваліфікаційних робіт рутинної роботою та розуміти цілеспрямоване та якісне використання інформації що видають LLM. Водночас, дасть навички оцінювати інформацію, яка видається у відповідь на запити на предмет адекватності та релевантності [7, 8]. Здобувачі зможуть відсіювати ті дані, що не відповідають дійсності та правильно інтерпретувати дані та факти, що видаються на запит від ГШІ.

Для коректного запровадження правил користування штучним інтелектом під час написання кваліфікаційних робіт необхідно провести аналіз та визначити характерні ознаки застосування таких інструментів в тексті та при організації структури та стилю написання. Такі референтні «відбитки» можуть допомагати в детекції неправомірного в академічному сенсі застосування ШІ.

Використовуючи власний досвід та досвід колег університету, можна зробити висновок, що здобувачі не розуміють, як коректно застосовувати штучний інтелект для оформлення власних результатів. Вони намагаються писати тексти та результати, які відповідають формальним вимогам, але по суті не можуть оцінити наскільки ця

інформація відповідає дійсності та науковому рівню. Це призводить до тексти кваліфікаційних робіт, хоч і виглядають як наукоємні, по факту не несуть в собі нічого нового і немає того необхідного додатково вкладу у власну кваліфікацію здобувача. Тому потрібно проводити лекції просвітницького характеру, навіть на спеціальностях недотичних до інформаційних технологій, щоб запобігти зловживанню та некоректному використанню такого складного та корисного інструменту як генеративний штучний інтелект.

Висновки. В роботі описано деякі з проблем, що виникають внаслідок некоректного використання ГШІ в кваліфікаційних роботах. Визначено декілька методів, для детекції та інтеграції інструментів штучного інтелекту в процес написання та оформлення кваліфікаційних робіт для здобувачів. В загальному випадку, визначення застосування ШІ зводиться до виявлення характерного стилю написання та відсутності так званого «людського фактору». Заборона застосування ШІ в певному сенсі є недоцільною, більш правильною стратегією є регламентування застосування та навчання здобувачів коректному застосуванню цих інструментів.

Література:

1. Sharples M. Automated Essay Writing: An AIED Opinion. *International Journal of Artificial Intelligence in Education*. 2022. URL: <https://doi.org/10.1007/s40593-022-00300-7> (date of access: 11.12.2024).
2. Sharples M. Towards social generative AI for education: theory, practices and ethics. *Learning: Research and Practice*. 2023. Vol. 9, no. 2. P. 159–167. URL: <https://doi.org/10.1080/23735082.2023.2261131> (date of access: 11.12.2024).
3. Sharples M. The Open University. *Open Research Online*. URL: <https://oro.open.ac.uk/78333/> (date of access: 11.12.2024).
4. Webb C., Luckin R., Ecoff E. Ethical Principles for the Development and Application of Artificial Intelligence in K-12 Education. *SpringerLink*. URL: https://link.springer.com/referenceworkentry/10.1007/978-981-19-6887-7_120 (date of access: 11.12.2024).
5. AI for Education (AI4EDU): Advancing Personalized Education with LLM and Adaptive Learning / Q. Wen et al. *KDD '24: The 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Barcelona Spain. New York, NY, USA, 2024. P. 6743–6744. URL: <https://doi.org/10.1145/3637528.3671498> (date of access: 11.12.2024).
6. Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin | Journal of Applied Learning and Teaching. *Public Knowledge Project*. URL: <https://journals.sfu.ca/jalt/index.php/jalt/article/view/1659> (date of access: 11.12.2024).
7. Palamar S., Naumenko M. Artificial Intelligence in Education: Use Without Violating the Principles of Academic Integrity. *Educological discourse*. 2024. Vol. 44, no. 1. P. 68–83. URL: <https://doi.org/10.28925/2312-5829.2024.15> (date of access: 11.12.2024).
8. ENAI Recommendations on the ethical use of Artificial Intelligence in Education / T. Foltynek et al. *International Journal for Educational Integrity*. 2023. Vol. 19, no. 1. URL: <https://doi.org/10.1007/s40979-023-00133-4> (date of access: 11.12.2024).