

Олег ЧЕРТОВ

доктор технічних, професор,
завідувач кафедри прикладної математики
Національного технічного університету України
«Київський політехнічний інститут ім. Ігоря Сікорського» (Київ)
chertov@i.ua

ЗАГРОЗИ ВИКОРИСТАННЯ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ І СУМНІВНИХ ПРАКТИК ДОСЛІДЖЕННЯ

У сучасному світі генеративний штучний інтелект (ГШІ) стає все більш популярним і широко використовується в різних галузях, від технічних до гуманітарних, від програмування до образотворчого мистецтва [1]. Однак, разом із зростанням його застосування не тільки з'являються нові можливості, але й виникають нові загрози, пов'язані з етичними питаннями, конфіденційністю, безпекою тощо. Ці загрози суттєво збільшуються, коли ГШІ використовується разом із сумнівними практиками дослідження (Questionable research practices, СПД).

Під сумнівними практиками дослідження в літературі розуміються певні методологічні, аналітичні або звітні практики, які викликають закономірні сумніви у їх застосовності, оскільки вони можуть бути використані для представлення упереджених доказів на користь певного твердження (і часто саме так й відбувається) [2]. Такі практики зазвичай призводять до неможливості відтворити або покращити опубліковані результати, тобто до так званої кризи відтворюваності (reproducibility crisis). Скажімо, відомий журнал Nature в 2016 році провів опитування 1576 дослідників з різних галузей науки і з'ясувалося, що:

- понад 70 % з них не змогли відтворити експерименти інших дослідників;

- більше половини не змогли відтворити результати своїх власних експериментів [3].

Наведені невтішні висновки ґрунтуються на порівнянні статистичної значущості результатів, отриманих в оригінальних дослідженнях, з результатами, отриманими в спробах їх повторення чи відтворення.

Існує велика кількість різноманітних сумнівних практик дослідження. Наприклад, в статті [4] виділено 34 ступені свободи, кожному з яких можна використовувати при експериментальних дослідженнях, при формулюванні гіпотез, плануванні, проведенні цих досліджень, а також при їх аналізі та звітності по ним. Тому лише перерахуємо ті сумнівні практики, які найчастіше

зустрічаються і при цьому мають напрочуд метафоричні назви [5]: просіювання даних (Data dredging), риболовля (Fishing trip), *P*-hacking, збирання стиглих вишень (Cherry-picking). Більш детальна інформація по цим практикам викладена в монографії [6], одним із співавторів якої є автор.

Застосування ГШІ додатково може ще зменшити довіру до наукових досліджень:

- (фальсифікація даних) використання згенерованих штучних інтелектом підроблених або маніпульованих даних для підтвердження гіпотез або отримання бажаних результатів призведе до недостовірних висновків, які тим не менше виглядатимуть як такі, що статистично підтверджені;

- (проблеми конфіденційності та безпеки) ГШІ може використовуватися для створення фальшивих ідентичностей або підробки даних, що може призвести до порушення конфіденційності та безпеки особистої інформації; особливо це актуально у випадках, коли ГШІ використовується для створення реалістичних зображень або відео;

- (плагіат) використання чужих ідей, методів або результатів штучним інтелектом без належного цитування порушує етичні норми та може призвести до втрати довіри до дослідника, який скористається ГШІ.

Тому серйозним викликом для всіх дослідників-експериментаторів є необхідність релевантного врахування в своїх роботах перерахованих загроз, викликаних спільним застосуванням ГШІ і СПД.

Література:

1. Feuerriegel S., Hartmann J., Janiesch C. and Zschech P. Generative AI. Business & Information Systems Engineering, vol. 66, pp. 111–126, 2024. DOI: 10.1007/s12599-023-00834-7
2. Banks G. C., O'Boyle E. H. Jr., Pollack J. M., White C. D., Batchelor J. H., Whelpley C. E., Abston K. A., Bennett A. A. and Adkins C. L. Questions about questionable research practices in the field of management: a guest commentary, Journal of Management, vol. 42, no. 1, pp. 5–20, 2016. DOI: 10.1177/0149206315619
3. Baker M. 1,500 scientists lift the lid on reproducibility, Nature, vol. 533, no. 7604, pp. 452–4, 26 May 2016. DOI: 10.1038/533452a
4. Wicherts J. M., Veldkamp C. L. S., Augusteijn H. E. M., Bakker M., van Aert R. C. M. and van Assen M. A. L. M. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking, Frontiers in Psychology, vol. 7, pp. 1–12, November 2016. DOI: 10.3389/fpsyg.2016.01832
5. Andrade C. HARKing, cherry-picking, P-hacking, fishing expeditions, and data dredging and mining as questionable research practices, Journal of Clinical Psychiatry, vol. 82, no. 1, 20f13804, 2021. DOI: 10.4088/JCP.20f13804
6. Demchenko, Y., Cuadrado-Gallego, J., Chertov, O., Aleksandrova, M. Big Data Infrastructure Technologies for Data Analytics: Scaling Data Science Applications for Continuous Growth. – Springer Cham, 2024. – XVI, 544 p.; ISBN: 9783031693656; DOI: 10.1007/978-3-031-69366-3