

ЕТИКА ЧАТ-СИСТЕМ ШІ В КОНТЕКСТІ КІБЕЗБЕЗПЕКИ: БАЛАНС МІЖ ПРИВАТНІСТЮ ТА ЗАХИСТОМ

Ситникова Ю.В.

Кафедра прикладної математики,
Харківський національний університет радіоелектроніки,
Україна.

E-mail: yuliia.sytnykova@nure.ua

Abstract

The ethical and cybersecurity implications of using an AI-based chat system are explored. Recently, the rapid advancement of language models has transformed human-machine interaction, but it has also increased the risks of data leakage, algorithmic bias, and social engineering. This study emphasises the importance of striking a balance between user privacy and information protection by applying the principles of privacy-by-design and ethics-by-design. Technical measures such as encryption, anonymisation of user data, and transparent data-handling policies are discussed. Moreover, ethical awareness among developers and users is emphasised as a critical prerequisite for responsible AI development. In conclusion, achieving a balance between privacy and security is crucial for ensuring trust, accountability, and the sustainable integration of AI chat technologies into the digital ecosystem.

У сучасному цифровому середовищі розвиток чат-систем на основі великих мовних моделей (Large Language Models, LLM) породжує низку нових ризиків у сфері кібербезпеки та захисту даних. Так, систематичні дослідження [1] зазначають, що чат-системи є дійсно вразливими до витоку даних, певних маніпуляцій та атак соціальної інженерії. Водночас українські наукові праці, присвячені дослідженню права на приватність у застосуванні ШІ [2] та аналізу потенціалу систем ШІ у забезпеченні кібербезпеки, підкреслюють актуальність проблеми етичного та безпечного використання таких технологій [3]. Попри значний науковий інтерес та численні наукові публікації, питання забезпечення балансу між приватністю та захистом інформації у чат-системах ШІ залишається недостатньо вивченим і потребує подальшого наукового аналізу й моніторингу, що визначає наукову й практичну значущість даної теми.

Забезпечення балансу між приватністю та захистом інформації (даних) має враховувати як технічний, так й етичний вимір. З одного боку, право на приватність є фундаментальною гарантією користувача в цифровому середовищі, а з іншого – надмірне його обмеження може призвести до значного зниження ефективності систем кібербезпечу, формування надлишкового контролю та, в окремих випадках, до цензури й дискримінації. Водночас науковці зазначають, що головною загрозою є не лише можливість витоку даних, але й непрозорість алгоритмів обробки інформації, складність відстеження процесів ухвалення рішення, а також потенційне зловживання з боку розробників або користувачів [2, 4]. Етичний аспект цієї проблеми полягає у визначенні меж допустимого збору, зберігання та аналізу даних користувачів, а також у формулюванні принципів відповідального використання ШІ, зокрема на засадах академічної доброчесності [5]. У зв'язку з цим подальше дослідження етичних засад функціонування чат-систем ШІ в контексті кібербезпеки набуває особливого значення. В цілому, загальні аспекти проблеми визначають рамки відповідальної поведінки всіх учасників цього процесу, починаючи з розробників й до користувачів, та включаючи фахівців з кібербезпеки. Все це створює певні передумови формування сучасного розуміння етичного використання технологій ШІ.

На сучасному етапі розвитку інформаційних технологій концептуальними засадами етики функціонування чат-систем ШІ є загальні принципи етики штучного інтелекту, які сформульовані в міжнародних документах [6, 7]. Вони передбачають пріоритет людської гідності, відкритість алгоритмів та захист приватності користувача. У контексті кібербезпеки ці засади доповнюються вимогою забезпечення надійності самих систем, запобігання маніпуляціям та недопущення дискримінаційних рішень. Таким чином, етика чат-систем ШІ може розглядатися як комплексна категорія, яка поєднує

технологічну безпеку з морально відповідальною поведінкою розробників, метою якої є формування довіри в цифровому середовищі. Інституціоналізація етичних принципів у технічні стандарти безпеки даних створює надійне підґрунтя для утвердження прозорості, справедливості, підзвітності та мінімізації загроз для користувачів. У межах сучасних підходів [8] важливою умовою етичного функціонування чат-систем є реалізація принципу *ethics-by-design* – інтеграція етичних норм безпосередньо на етапі проектування та розроблення технології ШІ. Проте на практиці реалізація цих засад стикається з низкою викликів, зокрема вразливість чат-систем щодо витоку даних, несанкціонованого впливу та упередженість алгоритмів. Ці ризики дійсно потребують подальшого аналізу з позиції кібербезпеки та вироблення практичних механізмів етичного захисту інформації.

Проведений аналіз етичних та кібербезпекових аспектів чат-систем ШІ підтверджує, що безпечне та відповідальне функціонування технологій можливе лише за умови комплексного підходу. Безсумнівно, ключовими викликами залишаються витоки даних, ризик маніпулювання користувачами та непрозорість алгоритмів, що підкреслює необхідність інтеграції механізмів захисту з етичними нормами поведінки, закріпленими в відповідних правових стандартах. На практиці це означає впровадження шифрування, анонімізації, контролю доступу, політик етичного використання даних, а також регулярне проведення аудит та моніторинг систем. До ефективних заходів слід віднести створення міждисциплінарних команд розробників, фахівців з кібербезпеки та етика для забезпечення прозорості, підзвітності та довіри користувачів. Все це, на нашу думку, має сприяти зниженню ризиків та підвищенню ефективності чат-систем у межах цифрової екосистеми.

Підсумовуючи вищевикладене, доцільно регулярно проводити аудит алгоритмів і моделей штучного інтелекту з метою виявлення упередженостей та вразливостей, а також забезпечувати навчання користувачів принципам безпечного використання чат-систем. Крім того, важливим є застосування міждисциплінарного підходу до розроблення алгоритмів функціонування таких систем із залученням фахівців з етики, права та кібербезпеки. У сукупності ці заходи націлено на комплексне управління ризиками та формування довіри в цифровому середовищі.

Література

1. Yang J., Chen Y.-L., Por L. Y., & Ku C. S. A Systematic Literature Review of Information Security in Chatbots. // *Applied Sciences*. 2023. Vol. 13, Is. 11. P. 63–55. Режим доступа: <https://doi.org/10.3390/app13116355>.
2. Попович Т.П., Маслюк О.В., Мацола А.А. Штучний інтелект та право на приватність: актуальні виклики цифрової епохи // *Науковий вісник Ужгородського національного університету. Серія: Право*. 2025. Т. 5, № 90. С. 465–470. Режим доступа: <https://doi.org/10.24144/2307-3322.2025.90.5.63>.
3. Даков С., Маньковський Д., Білоконь І. Системи штучного інтелекту в кібербезпеці та їхні можливості протистояти сучасним кіберзагрозам. // *Безпека інформаційних систем і технологій*. 2025. Т. 2, № 8. С. 42–48. Режим доступа: <https://doi.org/10.17721/ISTS.2024.8.42-48>.
4. Yao Y., Duan J., Xu K., Cai Y., Sun Z., Zhang Y. A survey on large language model (LLM) security and privacy: the good, the bad, and the ugly. // *High-Confidence Computing*. 2024. Vol. 4, Is. 2. Режим доступа: <https://doi.org/10.1016/j.hcc.2024.100211>.
5. Чудик Н. Етико-правові аспекти впровадження систем штучного інтелекту: балансування технологічного прогресу та фундаментальних прав людини в умовах цифрової трансформації суспільства. // *Актуальні проблеми правознавства*. 2024. №. 3. С. 42–47. Режим доступа: <https://appj.wunu.edu.ua/index.php/appj/article/view/1927>.
6. Recommendation on the ethics of Artificial Intelligence. UNESCO. 2021. [Електрон. ресурс] Режим доступа: <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>.
7. AI principles. 2019. [Електрон. ресурс] Режим доступа: <https://www.oecd.org/en/topics/ai-principles.html>
8. Brey P., Dainow B. Ethics by design for artificial intelligence. *AI Ethics*. 2024. Vol. 4. P. 1265–1277. Режим доступа: <https://doi.org/10.1007/s43681-023-00330-42024>.