Боровікова Віталіна Станіславівна
старший науковий співробітник,
Львівський державний університет внутрішніх справ
Scopus ID 58689328000
https://orcid.org/0000-0003-4401-4562

ЛАНЦЮГОВА ВІДПОВІДАЛЬНІСТЬ ЗА ДАНІ ЯК МЕХАНІЗМ УПРАВЛІННЯ РИЗИКАМИ ПРОТЯГОМ ЖИТТЄВОГО ЦИКЛУ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Стрімке впровадження технологій штучного інтелекту у сферу державного управління, охорони здоров'я, освіти, фінансів та цифрових сервісів актуалізує проблему належного управління даними протягом усього життєвого циклу алгоритмічних систем. Сучасні моделі ШІ є даноцентричними системами, ефективність, безпечність та справедливість яких безпосередньо залежать від якості даних, використаних для їх навчання, тестування та експлуатації. Як зазначають S. McGregor та J. Hostetler, саме управління даними поступово стає центральним елементом системи врядування ШІ, оскільки більшість ризиків виникає не лише через архітектуру алгоритмів, а й через характеристики даних, на яких вони функціонують [1].

У науковій літературі дедалі частіше підкреслюється необхідність переходу від оцінки окремих алгоритмів до комплексного управління ризиками протягом усього життєвого циклу систем штучного інтелекту. Такий підхід закріплено у міжнародних документах з врядування ШІ, зокрема в NIST AI Risk Management Framework, який передбачає безперервні процеси управління, оцінювання, моніторингу та контролю ризиків на етапах розроблення, впровадження та використання систем штучного інтелекту [2]. Аналогічний підхід простежується в європейській моделі регулювання, де особлива увага приділяється якості даних, простежуваності процесів та підзвітності учасників життєвого циклу високоризикових систем ШІ.

Особливого значення ця проблема набуває у контексті використання державних інформаційних ресурсів та електронних реєстрів як джерела даних для створення та навчання моделей штучного інтелекту. В умовах цифрової трансформації держави накопичуються значні масиви даних про громадян, юридичних осіб, адміністративні процедури та соціально-економічні процеси. Використання таких даних відкриває широкі можливості для автоматизації прийняття рішень, прогнозної аналітики та персоналізації публічних послуг.

Водночас виникають ризики порушення права на приватність, втрати контролю над персональними даними, алгоритмічної дискримінації та прийняття помилкових рішень на основі неповних або викривлених наборів даних.

Однією з головних проблем сучасного правового регулювання є відсутність чітко визначеного механізму розподілу відповідальності між усіма суб'єктами, які беруть участь у створенні та функціонуванні систем штучного інтелекту. На практиці відповідальність розподіляється між власниками даних, адміністраторами реєстрів, розробниками моделей, постачальниками цифрових сервісів та кінцевими користувачами. Проте у випадку виникнення шкоди або порушення прав людини часто складно встановити, на якому саме етапі життєвого циклу даних виникла помилка.

Для вирішення цієї проблеми доцільно застосовувати концепцію ланцюгової відповідальності за дані. Її сутність полягає у забезпеченні безперервної підзвітності всіх учасників процесу роботи з даними від моменту їх створення до завершення використання системи ШІ. У межах такого підходу кожен суб'єкт повинен нести відповідальність за якість, достовірність, законність походження та належність процедур обробки інформації на визначеному етапі життєвого циклу даних.

Концепція ланцюгової відповідальності узгоджується з сучасними принципами етичного та відповідального ШІ. У дослідженні L. Floridi та J. Cowls доведено, що ключовими засадами суспільно прийнятного використання штучного інтелекту є прозорість, справедливість, недопущення шкоди, автономія людини та підзвітність [3]. Реалізація цих принципів неможлива без належного контролю над даними та можливості відстеження їх походження і трансформації протягом усього життєвого циклу системи.

У практичному вимірі механізм ланцюгової відповідальності повинен включати документування походження даних (data provenance), аудит навчальних наборів, оцінювання ризиків повторної ідентифікації осіб, перевірку репрезентативності даних, моніторинг алгоритмічних упереджень та регулярну оцінку впливу систем ШІ на права людини. Важливим елементом такої моделі є створення журналів аудиту та забезпечення простежуваності рішень, які приймаються або підтримуються штучним інтелектом.

У сучасних дослідженнях наголошується, що ефективне врядування ШІ потребує переходу від декларативних принципів до конкретних організаційних процедур та механізмів контролю [4]. Саме тому управління даними повинно розглядатися як безперервний процес, інтегрований у всі етапи життєвого циклу

системи: проєктування, розроблення, навчання, тестування, впровадження, експлуатацію, моніторинг та виведення з експлуатації.

Для України формування механізму ланцюгової відповідальності за дані, наприклад, за допомогою дозвільної діяльності у сфері надання послуг електронних комунікацій [5, с. 88], є особливо актуальним з огляду на активну цифровізацію державного управління та перспективи використання технологій штучного інтелекту в діяльності органів влади. Впровадження такої моделі сприятиме підвищенню довіри до цифрових сервісів, забезпеченню правової визначеності, захисту прав людини та гармонізації національного законодавства з європейськими підходами до регулювання ШІ.

Отже, ланцюгова відповідальність за дані може розглядатися як один із базових інструментів побудови системи врядування штучного інтелекту, орієнтованої на принципи підзвітності, прозорості та управління ризиками. Подальший розвиток відповідних правових і організаційних механізмів дозволить забезпечити баланс між інноваційним розвитком технологій та належним захистом прав людини в умовах цифрового суспільства.

Список використаних джерел:

1. McGregor S., Hostetler J. Data-Centric Governance. arXiv, 2023. URL: https://arxiv.org/abs/2302.07872
2. Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.AI.100-1.
3. Floridi L., Cowls J. A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review. 2019. Vol. 1(1). DOI: 10.1162/99608f92.8cd550d1.
4. Dotan R., Blili-Hamelin B., Madhavan R., Matthews J., Scarpino J. Evolving AI Risk Management: A Maturity Model based on the NIST AI Risk Management Framework. arXiv, 2024. URL: https://arxiv.org/abs/2401.15229
5. Yesimov, S., & Borovikova, V. (2021). Licensing Activities in the Field of Provision of Telecommunications. *Social and Legal Studios*, 4(4), 83-89. https://doi.org/10.32518/2617-4162-2021-4-83-89