# Формування підходу до оцінювання ризиків використання штучного інтелекту в системі менеджменту інформаційної безпеки

УДК 004.056:004.8

Михайло Запорожченко[1], Світлана Легомінова[2], Дмитро Рабчун[3]

*Державний університет інформаційно-комунікаціних технологій,*
*[1]m.zaporozhchenko@duikt.edu.ua, [2]s.legominova@duikt.edu.ua,*
*[3]d.rabchun@duikt.edu.ua*

Інтенсивне впровадження систем штучного інтелекту (ШІ) в організаційні, управлінські та технологічні процеси зумовлює трансформацію умов забезпечення інформаційної безпеки (ІБ). Такі системи застосовуються для обробки даних, підготовки аналітичних матеріалів, автоматизації комунікацій, підтримки прийняття рішень, розробки програмного коду та виконання окремих функцій кіберзахисту. Водночас залучення ШІ-сервісів до роботи з корпоративною інформацією формує додаткові ризики, пов'язані з передачею чутливих даних до зовнішніх платформ, непрозорістю механізмів їх обробки, складністю перевірки згенерованих результатів, залежністю від постачальників і некритичним використанням автоматизованих висновків персоналом [1].

Проблема полягає в недостатній формалізації ризиків використання ШІ в процесах системи менеджменту інформаційної безпеки (СМІБ). У практичній діяльності такі ризики частково враховуються в політиках ІБ, вимогах до постачальників, процедурах управління інцидентами або правилах обробки даних, однак часто залишаються недостатньо інтегрованими в процеси ідентифікації, аналізу, обробки, моніторингу та перегляду ризиків. Внаслідок цього організація може встановлювати загальні обмеження щодо використання ШІ-сервісів, проте не мати методично визначеного механізму оцінювання рівня ризику для конкретних сценаріїв їх застосування.

Актуальність зазначеної проблеми зумовлена необхідністю включення ризиків використання ШІ до загальної логіки ризик-орієнтованого управління ІБ. Для організації важливим є не лише визначення допустимих і недопустимих способів застосування ШІ-сервісів, а й встановлення їхнього впливу на конфіденційність, цілісність, доступність, автентичність, підзвітність і контрольованість інформаційних процесів [2]. Особливого значення набуває оцінювання не тільки технічних характеристик ШІ-систем, а й організаційних умов їх використання, поведінки користувачів, вимог до обробки даних і можливості перевірки отриманих результатів.

Метою роботи є формування підходу до оцінювання ризиків використання ШІ в СМІБ шляхом визначення ризикоутворюючих чинників, параметрів інформаційного впливу та керованості відповідних сценаріїв застосування ШІ.

Для досягнення поставленої мети ризики використання ШІ запропоновано розглядати як сукупність чинників, що змінюють умови обробки, передачі, зберігання та подальшого використання інформації в організації. До них віднесено характер інформації, тип ШІ-сервісу та модель його розгортання, ступінь інтеграції з корпоративними системами, рівень автономності прийняття рішень, можливість перевірки результату, журналювання дій, компетентність

персоналу, вимоги постачальника щодо збереження даних і критичність процесу, у межах якого застосовується ШІ [3].

Запропонований підхід передбачає включення ризиків використання ШІ до чинного процесу управління ризиками в рамках СМІБ. Спочатку ідентифікуються сценарії застосування ШІ, залучені користувачі, процеси, інформаційні активи, типи даних і цілі використання відповідних систем. Далі кожен сценарій співвідноситься з потенційними наслідками для організації, зокрема з порушенням конфіденційності, цілісності, доступності, правових або договірних вимог, а також із впливом на управлінські, операційні чи технічні рішення. Оцінювання здійснюється за двома групами параметрів: інформаційного впливу та керованості ШІ-сценарію. Перша група характеризує чутливість даних, критичність процесу, масштаб можливих наслідків і значущість результатів ШІ для прийняття рішень; друга – наявність політик, технічних обмежень доступу, аудиту дій, перевірки постачальника, контролю умов збереження даних, вимог до валідації результатів і підготовки персоналу.

За результатами оцінювання визначаються заходи обробки ризику, рівень жорсткості яких має відповідати критичності сценарію застосування ШІ. Для низькоризикових сценаріїв достатніми можуть бути загальні правила використання та інформування персоналу; для середньоризикових – перелік дозволених сервісів, обмеження щодо категорій даних і вимоги до перевірки результатів; для високоризикових – використання корпоративно контрольованих ШІ-рішень, журналювання дій, договірні вимоги до постачальників, регулярний перегляд ризиків і включення відповідних сценаріїв до програми внутрішнього аудиту.

Запропонований підхід дозволяє конкретизувати місце ризиків використання ШІ в структурі СМІБ, узгодити їх оцінювання з процесами управління ризиками ІБ, і пов'язати сценарії використання ШІ з контекстом організації, вимогами зацікавлених сторін, інформаційними активами, процедурами обробки ризиків, моніторингом результативності заходів захисту та внутрішнім аудитом. При цьому акцент переноситься з формального дозволу або заборони ШІ-сервісів на визначення умов, за яких їх використання може бути прийнятним, обмеженим або недопустимим з погляду ІБ.

За результатами дослідження обґрунтовано доцільність розгляду ризиків використання ШІ як чинника, що змінює умови обробки інформації, межі відповідальності, вимоги до контролю результатів і рівень керованості інформаційних процесів у межах СМІБ. Основним результатом є уточнення логіки оцінювання таких ризиків через поєднання параметрів інформаційного впливу та параметрів керованості сценарію застосування ШІ.

1. Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.AI.100-1.
2. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection – Information security management systems – Requirements. International Organization for Standardization, 2022.
3. ISO/IEC 42001:2023. Information technology – Artificial intelligence – Management system. International Organization for Standardization, 2023.