УДК 004.8:303.725.36
DOI https://doi.org/10.32782/2786-5681-2025-1.15

Ольга КРАВЧУК
кандидат політичних наук, доцент кафедри соціально-гуманітарних наук та філософії, Національний університет кораблебудування імені адмірала Макарова
olha.kravchuk@nuos.edu.ua
ORCID: *0000-0001-7802-1934*

ПРОЦЕДУРА ІДЕНТИФІКАЦІЇ РИЗИКІВ ВПРОВАДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ПУБЛІЧНЕ УПРАВЛІННЯ

Анотація. *Ідентифікація ризиків впровадження штучного інтелекту (ШІ) у публічне управління є ключовим етапом забезпечення ефективності та безпеки цього процесу. Ця процедура передбачає систематичний аналіз потенційних загроз, які можуть виникнути під час інтеграції технологій ШІ в управлінські процеси, включаючи прийняття рішень, надання послуг та управління даними. Першим етапом є визначення сфер застосування ШІ, таких як автоматизація рутинних процесів та аналіз великих обсягів даних. Далі слід оцінити технічні, етичні, соціальні та правові ризики. Технічні ризики можуть включати помилки в алгоритмах та збої у системах, тоді як етичні ризики охоплюють проблеми конфіденційності даних і упередженість алгоритмів. Соціальні ризики можуть проявлятися у зниженні довіри громадян до публічного управління, а правові ризики стосуються невідповідності впровадження ШІ чинному законодавству. Для ідентифікації ризиків доцільно використовувати методи, такі як SWOT-аналіз та експертне опитування. Результати аналізу ризиків повинні бути задокументовані та слугувати основою для розробки стратегій їх мінімізації, зокрема шляхом впровадження заходів кібербезпеки та створення етичних стандартів.*

Метою статті *є вивчення сучасних теоретичних підходів до оцінки ризиків та пропозиція моделі процедури ідентифікації ризиків впровадження ШІ в публічне управління.*

Методологія дослідження *передбачає застосування комбінації якісних і кількісних методів, включаючи аналіз літератури та експертне опитування.*

Наукова новизна *роботи полягає у систематичному аналізі ризиків, пов'язаних із впровадженням ШІ, та розробці нових підходів до їх управління.* ***Висновки****. Заходи управління ризиками вимагають регулярного перегляду для адаптації до нових умов і технологій. Запропонована модель процедури ідентифікації ризиків дозволяє ефективно визначати та мінімізувати потенційні загрози, що сприятиме зростанню довіри до державних інститутів і підвищить ефективність публічних послуг.*

Ключові слова: *публічне управління, штучний інтелект, ризики, види ризиків, інтеграція технологій, етичні стандарти, загрози.*

Olha KRAVCHUK
Candidate of Political Science, Associate Professor at the Department of Social, Humanitarian and Philosophy, Admiral Makarov National University of Shipbuilding
olha.kravchuk@nuos.edu.ua
ORCID: *0000-0001-7802-1934*

THE PROCEDURE FOR IDENTIFYING THE RISKS OF INTRODUCING ARTIFICIAL INTELLIGENCE INTO PUBLIC ADMINISTRATION

Abstract. *Identifying the risks of implementing artificial intelligence (AI) in public administration is a key step in ensuring the effectiveness and security of this process. This procedure involves a systematic analysis of potential threats that may arise when integrating AI technologies into administrative processes, including decision-making, service provision and data management. The first stage is to identify areas of application of AI, such as automation of routine processes and analysis of large volumes of data. Next, technical, ethical, social and legal risks should be assessed. Technical risks may include errors in algorithms and system failures, while ethical risks include data confidentiality issues and algorithm bias. Social risks may manifest themselves in reducing citizens' trust in public administration, and legal risks relate to the inconsistency of AI implementation with applicable legislation. To identify risks, it is advisable to use methods such as SWOT analysis and expert survey.*

The aim of the article *is to study modern theoretical approaches to risk assessment and propose a model for the procedure for identifying risks of implementing AI in public administration.*

The research methodology *involves the use of a combination of qualitative and quantitative methods, including literature analysis and expert survey.*

The scientific novelty *of the work lies in the systematic analysis of risks associated with the implementation of AI and the development of new approaches to their management. Risk management measures require regular revision to adapt to new conditions and technologies.*

Conclusions*. The proposed model of the risk identification procedure allows for the effective identification and minimization of potential threats, which will contribute to increasing trust in state institutions and increasing the efficiency of public services.*

Key words: *public administration, artificial intelligence, risks, types of risks, integration of technologies, ethical standards, threats.*

Постановка проблеми. Впровадження штучного інтелекту (ШІ) в публічне управління відкриває значні можливості для підвищення ефективності державних послуг, автоматизації процесів, прогнозування соціальних потреб та прийняття більш обґрунтованих управлінських рішень. Однак одночасно це створює нові ризики, які можуть мати серйозні наслідки для суспільства, прав і приватності громадян, а також для стабільності й репутації державних органів. Серед них можемо назвати вразливість до технологічних ризиків і кіберзагроз, ризики порушення приватності та конфіденційності, соціальні ризики та недовіра громадян, недостатній рівень компетенцій посадових осіб.

Ідентифікація потенційних вигод від впровадження штучного інтелекту (ШІ) в публічне управління є важливим етапом, що допомагає державним органам та громадянам зрозуміти основні цілі й переваги використання цієї технології. Для державних органів це означає підвищення ефективності роботи завдяки автоматизації рутинних процесів, що дозволяє скоротити витрати часу та ресурсів, а також зменшити кількість помилок у виконанні стандартних операцій. ШІ може допомогти в обробці великих обсягів даних, швидкому отриманні аналітичної інформації та генерації прогнозів, що сприяє прийняттю більш обґрунтованих та стратегічних управлінських рішень. Зі свого боку, громадяни можуть очікувати покращення якості державних послуг, оскільки автоматизовані процеси дозволяють значно скоротити час обслуговування та підвищити його точність. Додатково, за допомогою аналітичних можливостей ШІ державні органи здатні краще адаптувати послуги під потреби населення, що позитивно впливає на загальне задоволення громадян.

Визначення меж і сфери застосування ШІ є важливим для того, щоб уникнути непорозумінь і завищених очікувань. Зокрема, ШІ можна використовувати для автоматизації рутинних процесів, таких як обробка заяв, заявок та інших типових процедур, що вимагають стандартних дій. У сфері аналітики ШІ допомагає обробляти та аналізувати дані в реальному часі, що є корисним для планування, оцінки ефективності та прийняття рішень на основі фактичних даних. Крім того, ШІ може підтримувати процеси прийняття рішень, надаючи державним службовцям розгорнуті прогнози, рекомендації або аналітичні звіти. Однак важливо також обмежити сферу дій ШІ, залишаючи остаточне рішення за людиною, щоб забезпечити відповідальність і врахування соціальних аспектів, які ШІ може не повністю зрозуміти.

Аналіз джерел та останніх досліджень. Марутян Р.Р. досліджує практики використання новітніх інформаційних технологій інтелектуального управління (блокчейн, штучний інтелект, Інтернет речей тощо) у публічному управлінні різних країн світу та в Україні; охарактеризовано їх важливість для впровадження e-Government та ризики їх застосування у сучасній публічно-управлінській практиці [8].

Етичні питання є ще одним важливим напрямом досліджень. Праці таких авторів, як А. Флоріді та В. Хесс, зосереджені на розгляді етичних викликів, які виникають при впровадженні ШІ в публічному управлінні, включаючи питання справедливості, прозорості й підзвітності. Вони відзначають, що державні органи мають додержуватись високих етичних стандартів, і закликають до створення «етичних рад» або наглядових комітетів, які б забезпечували контроль за дотриманням принципів етики в роботі ШІ [4].

Деякі автори також підкреслюють проблему довіри до рішень ШІ, оскільки люди можуть сумніватися в об’єктивності алгоритмів. Дослідження Л. Корнута показує, що для підвищення рівня довіри громадськості важливо забезпечити прозорість процесу прийняття рішень і надати можливість громадянам розуміти, як працює ШІ [1].

Метою даної статті є вивчення сучасних теоретичних підходів та практичних методик

до оцінки ризиків, пропозиція моделі процедури ідентифікації ризиків впровадження ШІ в публічне управління.

Виклад основного матеріалу. Процедура ідентифікації ризиків впровадження штучного інтелекту (ШІ) в публічне управління має ряд особливостей, пов'язаних з високим рівнем складності, міждисциплінарністю та різноманітністю сфер впливу ШІ. Перш за все, ШІ зачіпає різні рівні та сфери державного управління, включаючи політику, правові аспекти, управлінські процеси та технологічну інфраструктуру. Тому важливо, щоб процедура враховувала можливий системний вплив ШІ на численні державні установи та процеси, аналізуючи ризики в комплексі. Процедура повинна бути міждисциплінарною, залучаючи до оцінки експертів з різних галузей – технічних спеціалістів, юристів, етиків, фахівців з державного управління та представників громадськості. Це дозволяє виявити як технічні, так і соціальні, етичні та правові ризики, пов'язані з використанням ШІ.

Особливу увагу варто приділити потенційним загрозам для безпеки, приватності громадян, прозорості та справедливості прийняття рішень. Наприклад, внаслідок використання ШІ можуть з'являтися загрози конфіденційності персональних даних, упередженості алгоритмів, непрозорості результатів і рішень, які він генерує. Також необхідно передбачити можливість появи технологічних збоїв і кіберзагроз, адже під час інтеграції нових технологій в державне управління важливо зберегти довіру суспільства. Процедура ідентифікації ризиків має включати аналіз потенційних впливів на права людини, вплив на організаційні процеси в державних установах і можливість несправедливого або дискримінаційного впливу ШІ на різні соціальні групи [2]. Програми штучного інтелекту можуть потенційно підвищити ефективність і результативність надання адміністративних послуг і підтримувати прийняття урядових рішень шляхом імітації різних варіантів політики при виконанні функцій публічного управління. Дослідники вказали на потенціал покращення розробки політики за допомогою технологій штучного інтелекту, надаючи державним службовцям додаткову інформацію, отриману на основі даних, автоматизувати повсякденні завдання та процеси, покращити інформацію, що надається громадянам, зробити послуги більш персоналізованими і краще зрозуміти настрої та потреби громадян, наприклад, за допомогою аналізу даних із соціальних мереж. Однак більшість потенційних наслідків ШІ (позитивних чи негативних) ще належить підтвердити й оцінити на емпіричній основі. Це пов'язано з різними бар'єрами, з якими громадські організації стикаються в інноваціях, які включають кілька факторів, що походять від середовища, організаційного контексту, рівня інновацій, а також індивідуальних пов'язаних факторів, до певної міри унікальних у контексті державного сектора порівняно з приватним сектором [4].

Алгоритм ідентифікації ризиків впровадження штучного інтелекту (ШІ) в публічне управління складається з кількох послідовних етапів, що дозволяють системно виявити потенційні загрози, оцінити їх значущість і підготувати відповідні заходи для мінімізації негативного впливу.

Перший етап – визначення цілей і завдань впровадження ШІ. На цьому етапі формується чітке розуміння, навіщо ШІ впроваджується, яких результатів очікують досягти, а також які конкретні процеси чи завдання він буде підтримувати чи автоматизувати. Розуміння кінцевих цілей допомагає визначити ризики, пов'язані з конкретними функціями та діями, які виконує ШІ.

Другий етап – ідентифікація категорій ризиків. Ризики можуть бути різного типу: операційні, соціальні, правові, політичні, економічні та технологічні. Наприклад, операційні ризики пов'язані з можливими помилками в алгоритмах, соціальні ризики можуть стосуватися дискримінації або упередженості алгоритмів, а технологічні ризики включають загрози кібербезпеці та обмеження самої технології. Визначення категорій ризиків допомагає систематизувати загрози та забезпечити, щоб жоден тип потенційної небезпеки не залишився поза увагою.

Третій етап – аналіз сценаріїв можливого використання ШІ. Для кожного сценарію використання (наприклад, автоматизація документообігу або підтримка аналітичних процесів) аналізуються відповідні ризики. Це дозволяє врахувати специфіку кожного випадку, оскільки ризики можуть суттєво відрізнятися залежно від того, як саме буде застосовано ШІ.

Четвертий етап – визначення джерел ризиків. На цьому етапі ідентифікуються конкретні джерела загроз, як-от людські помилки при проєктуванні алгоритмів, недосконаліть вихідних даних, технічні обмеження ШІ-системи або зовнішні фактори, наприклад, зміна регуляторної політики. Визначення джерел ризиків дозволяє не тільки краще їх зрозуміти, а й визначити можливі способи запобігання цим ризикам.

П'ятий етап – оцінка ймовірності та серйозності кожного ризику. Для цього ризики оцінюють за шкалою ймовірності (високий, середній, низький) та шкалою серйозності їхнього впливу на державну установу та суспільство загалом. Це допомагає зрозуміти, які ризики потребують першочергової уваги, а які є менш критичними.

Шостий етап – пріоритезація ризиків на основі оцінок. Після оцінки ймовірності та серйозності формують перелік ризиків за пріоритетом, де найвищий пріоритет мають ризики з високою ймовірністю та серйозністю. Така пріоритезація дозволяє зосередити ресурси на найбільш загрозливих аспектах впровадження ШІ.

Сьомий етап – розробка попереджувальних заходів для мінімізації або уникнення ризиків. Це можуть бути заходи, спрямовані на поліпшення якості алгоритмів, усунення упередженості даних, посилення кіберзахисту чи забезпечення відповідності регуляторним вимогам.

Розробка заходів управління ризиками є важливим етапом для забезпечення надійного впровадження штучного інтелекту (ШІ) в публічне управління. Цей процес розпочинається з визначення попереджувальних заходів, які дозволяють уникнути або мінімізувати ймовірність настання ідентифікованих ризиків ще до їх реалізації. Наприклад, для уникнення технічних збоїв і упередженості алгоритмів проводиться ретельне тестування системи ШІ, зокрема в умовах, що максимально наближені до реальних, а також періодична оцінка алгоритмів на предмет упередженості та відповідності нормативним вимогам. Окремо важливо забезпечити відповідність системи правовим стандартам, що передбачає регулярний моніторинг змін у законодавстві, пов'язаному з конфіденційністю даних та захистом прав громадян.

Наступним кроком є розробка мігітаційних заходів, які дозволяють зменшити негативний вплив ризиків, що вже виникли. Наприклад, створення резервних систем, які можуть взяти на себе функції ШІ в разі технічного збою, або розробка плану на випадок кіберінциденту для швидкого відновлення нормальної роботи системи. Інший важливий аспект мігітаційних заходів – це навчання співробітників правильній роботі з системами ШІ, щоб мінімізувати ризик людської помилки.

Стратегія реагування на ризики включає створення чітких процедур для швидкого виявлення та управління ризиками у режимі реального часу. Для цього варто впровадити системи моніторингу та відстеження, які б автоматично виявляли відхилення у функціонуванні ШІ або інші сигнали про ризик. Важливим є також план дій на випадок надзвичайних ситуацій, що включає кроки для зменшення наслідків ризику та його нейтралізації. Наприклад, при виявленні помилкових рішень ШІ повинні бути передбачені процедури для швидкого перегляду таких рішень або їх відміни з можливістю подальшої корекції.

Висновки. Заходи управління ризиками вимагають регулярного перегляду та оновлення для адаптації до нових умов, зміни технологій та правових норм. Це передбачає також проведення аудитів та оцінок ризиків після впровадження ШІ, щоб перевірити ефективність реалізованих заходів, а також внесення корективів на основі отриманих даних та зворотного зв'язку від користувачів і працівників.

Узагальнюючи дослідження проблеми ідентифікації ризиків впровадження штучного інтелекту в публічне управління, можна зробити висновок, що інтеграція технологій ШІ у державні процеси має значний потенціал, але також супроводжується численними ризиками, які потребують системного підходу до їхньої ідентифікації та управління. Запропонована модель процедури ідентифікації ризиків, яка включає етапи моніторингу, оцінки та управління, дає змогу ефективно визначати та мінімізувати потенційні загрози на всіх стадіях використання ШІ, від збору та обробки даних до прийняття управлінських рішень.

Наявність надійної процедури управління ризиками сприятиме зростанню довіри до державних інститутів, забезпечить захист прав громадян і підвищить ефективність публічних послуг. Важливими аспектами є також дотримання правових і етичних стандартів, що запобігають дискримінації та порушенню приват-

ності громадян. Це підкреслює необхідність комплексного підходу, який би об'єднував технологічні, соціальні, правові та етичні компоненти. Подальші дослідження і вдосконалення механізмів моніторингу та оцінки ризиків є перспективним напрямом, що забезпечить безпечне і відповідальне впровадження ШІ в державному секторі.

ЛІТЕРАТУРА:

1. Корнута Л. Штучний інтелект у публічному управлінні: перспективи впровадження. *Європейські орієнтири розвитку України в умовах війни та глобальних викликів XXI століття: синергія наукових, освітніх та технологічних рішень*. Національний університет «Одеська юридична академія». Матеріали Міжнародної науково-практичної конференції 19 травня 2023 року.

2. Максименцева Н., Максименцев М. Штучний інтелект у публічному управлінні: переваги цифрових технологій та загрози суверенному інформаційному простору. *Державне управління: удосконалення та розвиток*. 2024. № 2.

3. Марутян Р.Р. Інформаційні технології інтелектуального управління у публічно-управлінській практиці: зарубіжний та вітчизняний досвід. *Вісник Національного університету цивільного захисту України*. 2018. № 2. С. 146–153.

4. Artificial Intelligence (AI) In Cyber Security Market Will Reach to USD 30.9 Billion By 2025: Zion Market Research. URL: https://www.globenewswire.com/news-release/2019/08/28/1907655/0/en/Artificial-Intelligence-AI-In-Cyber-Security-Market-Will-Reach-to-USD-30-9-Billion-By-2025-Zion-Market-Research.html.

5. Artificial Intelligence in Cybersecurity Market by Offering. URL: https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-security-market-220634996.html.

6. Eriksson M., Djoweini C. Artificial Intelligence's Impact on Management: A literature review covering artificial intelligence's influence on leadership skills and managerial decision-making processes : thesis. 2020. URL: http://urn.kb.se/resolveurn=urn:nbn:se:kth:diva-279737

7. Ethics guidelines for trustworthy AI. URL: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

8. Leveraging artificial intelligence to maximize critical infrastructure cybersecurity. URL: https://www.thalesgroup.com/en/worldwide/security/magazine/leveraging-artificial-intelligence-maximize-criticalinfrastructure.

REFERENCES:

1. Kornuta, L. (2023). Shtuchnyj intelekt u publichnomu upravlinni: perspektyvy vprovadzhennya. [Artificial Intelligence in Public Administration: Implementation Prospects]. *Yevropejski oriyentyryrozvytku ukrayiny v umovax vijnyta globalnyx vyklykiv XXI stolittya: synergiya naukovyx, osvitnix ta texnologichnyx rishen`*. Nacionalnyj universytet «Odeska yurydychna akademiya» [in Ukrainian].

2. Maksymenceva, N., & Maksymencev, M. (2024). Shtuchnyj intelekt u publichnomu upravlinni: perevagycyfrovyx texnologij ta zagrozy suverennomu informacijnomu prostoru [Artificial Intelligence in Public Administration: Benefits of Digital Technologies and Threats to Sovereign Information Space]. *Derzhavne upravlinnia: udoskonalennia ta rozvytok*, 2 [in Ukrainian].

3. Marutyan, R.R. (2018). Informacijni texnologiyi intelektualnogo upravlinnya u publichno-upravlinskij praktyci: zarubizhnyj ta vitchyznyanyj dosvid. [Information technologies of intellectual management in public management practice: foreign and domestic experience]. *Visnyk Natsionalnoho universytetu tsyvilnoho zakhystu Ukrainy,* 2, 146–153 [in Ukrainian].

4. Artificial Intelligence (AI) In Cyber Security Market Will Reach to USD 30.9 Billion By 2025: Zion Market Research. Retrieved from: https://www.globenewswire.com/news-release/2019/08/28/1907655/0/en/Artificial-Intelligence-AI-In-Cyber-Security-Market-Will-Reach-to-USD-30-9-Billion-By-2025-Zion-Market-Research.html [in English].

5. Artificial Intelligence in Cybersecurity Market by Offering. Retrieved from: https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-security-market-220634996.html [in English].

6. Eriksson, M., & Djoweini, C. (2020). [Artificial Intelligence's Impact on Management: A literature review covering artificial intelligence's influence on leadership skills and managerial decision-making processes] : thesis. 2020. Retrieved from: http://urn.kb.se/resolveurn=urn:nbn:se:kth:diva-279737

7. Ethics guidelines for trustworthy AI. Retrieved from: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai [in English].

8. Leveraging artificial intelligence to maximize critical infrastructure cybersecurity. Retrieved from: https://www.thalesgroup.com/en/worldwide/security/magazine/leveraging-artificial-intelligence-maximize-criticalinfrastructure [in English].