УДК 004.89
DOI https://doi.org/10.32782/TNU-2663-6468/2024.4/13

Кравчук О.Ю.
Національний університет кораблебудування імені адмірала Макарова

ЗАБЕЗПЕЧЕННЯ ЕТИЧНИХ СТАНДАРТІВ ТА БЕЗПЕКИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У ПУБЛІЧНОМУ УПРАВЛІННІ

У статті розглядаються ключові аспекти забезпечення етичних стандартів і безпеки при використанні ШІ в публічному управлінні, включаючи питання конфіденційності, прозорості, відповідальності та уникнення дискримінації. Автори також аналізують міжнародний досвід і наводять рекомендації для впровадження найкращих практик у цій сфері. З розвитком технологій штучного інтелекту (ШІ) та його інтеграцією в різні сфери суспільного життя, включаючи публічне управління, виникають нові етичні та безпекові виклики. У сучасному світі, де штучний інтелект (ШІ) стає невід'ємною частиною суспільного життя, його впровадження у сферу публічного управління викликає значний інтерес. Однак разом із потенційними перевагами, які ШІ може принести для підвищення ефективності управління, існує низка викликів, пов'язаних з етикою та безпекою його використання. Дана стаття присвячена аналізу основних етичних стандартів та безпекових аспектів, що повинні бути враховані під час інтеграції ШІ в публічне управління. Автор досліджує правові, соціальні та технічні питання, які виникають при використанні ШІ в державних структурах. Особлива увага приділяється питанням конфіденційності та захисту персональних даних, що є надзвичайно важливими в контексті державного управління. Окрім цього, у статті розглядаються проблеми відповідальності та прозорості в прийнятті рішень, заснованих на алгоритмах ШІ, що може впливати на права та свободи громадян. У статті також аналізуються міжнародні практики та рекомендації щодо забезпечення етичного використання ШІ, включаючи розробку відповідних нормативно-правових актів та створення спеціальних регуляторних органів. Автори підкреслюють необхідність розробки етичних кодексів і стандартів для публічного управління з урахуванням специфіки застосування ШІ, а також важливість підвищення обізнаності державних службовців та широкої громадськості щодо потенційних ризиків, пов'язаних із ШІ. На основі проведеного аналізу у статті пропонуються рекомендації щодо впровадження комплексних підходів до забезпечення етичних стандартів і безпеки ШІ, які включають технічні, організаційні та нормативно-правові заходи. Висновки статті спрямовані на сприяння формуванню відповідального підходу до використання ШІ в публічному управлінні, що дозволить максимізувати його позитивний вплив і мінімізувати можливі негативні наслідки.

Ключові слова: *штучний інтелект, публічне управління, етика, безпека, конфіденційність, прозорість, публічне управління.*

Постановка проблеми. Використання штучного інтелекту в публічному управлінні стає дедалі поширенішим явищем, завдяки якому уряди можуть підвищити ефективність надання послуг, оптимізувати процеси ухвалення рішень і забезпечити більш інклюзивне управління. Однак разом із цими перевагами виникають серйозні етичні та безпекові проблеми, які необхідно вирішувати для збереження довіри громадськості до державних інституцій.

Одними з основних етичних питань використання ШІ в публічному управлінні є прозорість і підзвітність. Прозорість алгоритмів ШІ є ключовим фактором, який визначає довіру громадян до системи публічного управління. Відкритість та можливість аудиту рішень, прийнятих на основі ШІ, дозволяє запобігти зловживанням і підвищує рівень підзвітності посадових осіб. Прозорість та підзвітність є двома ключовими аспектами етичного використання штучного інтелекту (ШІ) у публічному управлінні. Вони визначають, наскільки громадяни та інші зацікавлені сторони можуть довіряти системам, що базуються на ШІ, і наскільки ці системи відповідають за свої дії та рішення.

Аналіз останніх досліджень і публікацій. Ряд дослідників виступають за більш активне впровадження технологій ШІ в роботу урядів. Так, О. Карпенко докладно аналізує історію ство-

рення та розвитку ШІ, представлено основні сучасні підходи до визначення поняття, показано шляхи та прогнози подальшого використання ШІ в різних сферах життєдіяльності суспільства та публічного управління. О. Оболенський досліджує перспективи використання ШІ в публічному управлінні та розглядає етичні аспекти цього застосування. Результати дослідження підкреслюють використання ШІ для покращення ефективності, прозорості та інноваційності в системі публічного управління. Ця тема є водночас дуже актуальною та майже нерозкритою в Україні. [2]

Постановка завдання. Метою статті є дослідження основних етичних викликів, пов'язаних з впровадженням штучного інтелекту у публічному управлінні, включаючи питання справедливості, прозорості, приватності та відповідальності. Визначити потенційні загрози для безпеки, які можуть виникнути при використанні ШІ у сфері публічного управління, зокрема ризики зловживання даними, кібератак та порушення конфіденційності. Дослідити існуючі правові рамки та стандарти регулювання використання ШІ у публічному управлінні, а також запропонувати напрямки для їх удосконалення.

Виклад основного матеріалу. Велику роль відіграє такий момент, як те, що інтерфейси, через які громадяни взаємодіють з системами ШІ, мають бути простими та інтуїтивно зрозумілими. Це необхідно для того, щоб користувачі могли повноцінно розуміти процеси, які відбуваються у системі, і впливати на них у межах своїх повноважень.

У випадках, коли ШІ приймає рішення, важливо чітко визначити, хто несе відповідальність за ці рішення – чи то розробники, чи користувачі системи, чи державні органи, які її впроваджують. Це допомагає запобігти безвідповідальності та забезпечити належне реагування на можливі помилки або зловживання.

Громадяни мають право оскаржувати рішення, прийняті на основі ШІ, якщо вони вважають їх несправедливими або необґрунтованими. Для цього необхідно створити відповідні правові та процедурні механізми, які дозволяють швидко та ефективно розглядати такі скарги.

Не треба забувати, що підзвітність передбачає регулярний аудит та моніторинг роботи систем ШІ. Це можуть бути як внутрішні перевірки, так і незалежний зовнішній аудит. Метою цих заходів є перевірка відповідності роботи систем ШІ етичним стандартам, законодавчим вимогам та суспільним очікуванням.

Органи публічного управління, що використовують ШІ, мають регулярно звітувати перед громадськістю про те, як використовуються ці технології, які рішення приймаються на їх основі, та які заходи вживаються для забезпечення прозорості та підзвітності. Такі звіти сприяють підвищенню довіри громадян до державних інституцій та стимулюють розвиток етичних стандартів у сфері ШІ.

Забезпечення прозорості та підзвітності ШІ у публічному управлінні стикається з низкою викликів, включаючи складність пояснення складних алгоритмів, ризик порушення конфіденційності під час відкриття даних та можливі технічні обмеження. Проте ці виклики можуть бути подолані через розробку більш ефективних технологічних рішень, правових регуляцій та етичних стандартів.

Отже, прозорість та підзвітність є фундаментальними принципами, які мають бути дотримані при використанні ШІ в публічному управлінні. Їх забезпечення є запорукою підвищення довіри громадськості, ефективності державних послуг та відповідальності органів влади. Для цього необхідно розробляти та впроваджувати відповідні правові рамки, технологічні рішення та практики, які забезпечуватимуть дотримання цих принципів.

Наступним важливим етичним питанням в процесі застосування технологій та ШІ виступає конфіденційність та захист персональних даних. Воно постає здебільшого у зв'язку з тим, що ШІ-системи в публічному управлінні часто працюють з великим обсягом персональних даних, що піднімає питання конфіденційності. Важливо забезпечити захист цих даних від несанкціонованого доступу та використання, що вимагає застосування відповідних технологічних та правових заходів.

Застосування штучного інтелекту (ШІ) в публічному управлінні сприяє підвищенню ефективності державних послуг, однак це також породжує значні виклики, пов'язані з конфіденційністю та захистом персональних даних. Забезпечення безпеки даних є критичним аспектом, оскільки державні установи часто обробляють великі обсяги чутливої інформації про громадян. Тож доречно буде розглянути власне конфіденційність у контексті ШІ.

1. Обробка персональних даних. ШІ-системи в публічному управлінні часто використовують персональні дані громадян для аналізу, прогнозування та прийняття рішень. Це можуть бути дані про здоров'я, фінансовий стан, соціальний статус тощо. Використання цих даних вимагає дотримання принципів конфіденційності, які включа-

ють мінімізацію збору даних, використання їх лише для визначених цілей і забезпечення доступу до них лише уповноваженим особам.

2. Анонімізація та псевдонімізація Щоб знизити ризик ідентифікації особи на основі персональних даних, використовуються методи анонімізації та псевдонімізації. Анонімізація передбачає видалення або модифікацію даних так, щоб особу не можна було ідентифікувати, тоді як псевдонімізація замінює ідентифікаційні дані на псевдоніми, які можуть бути розшифровані лише за наявності спеціального ключа.

3. Консенсус та інформованість громадян. Прозорість у зборі та обробці персональних даних передбачає, що громадяни повинні бути поінформовані про те, які дані збираються, з якою метою і як вони будуть використовуватись. Важливо забезпечити, щоб збір даних здійснювався на основі згоди громадян або в рамках законних повноважень, а також щоб громадяни мали доступ до своїх даних і могли контролювати їх використання.

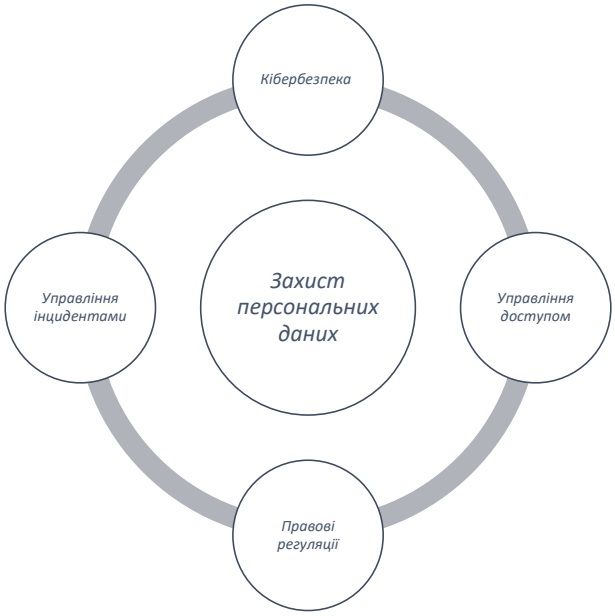

Рис. 1

1. Кібербезпека. ШІ-системи повинні бути захищені від кіберзагроз, таких як несанкціонований доступ, викрадення або пошкодження даних. Це вимагає застосування сучасних технологій шифрування, систем виявлення вторгнень, багатофакторної автентифікації та інших засобів захисту. Важливо також проводити регулярний аудит безпеки систем та забезпечувати їх оновлення.

2. Управління доступом. Необхідно обмежувати доступ до персональних даних виключно тими співробітниками, які мають відповідні повноваження та необхідність у доступі для виконання своїх обов'язків. Це передбачає впровадження систем управління доступом і контроль за використанням даних.

3. Правові регуляції. Для захисту персональних даних на рівні законодавства важливо впроваджувати та дотримуватися відповідних регуляцій, таких як Загальний регламент захисту даних (GDPR) у Європейському Союзі або інші національні законодавства, що регулюють обробку персональних даних. Це дозволяє встановити чіткі правила для державних установ, які використовують ШІ, та накладати санкції за їх порушення.

4. Управління інцидентами. У випадку порушення конфіденційності або витоку даних, важливо мати чіткі процедури для швидкого реагування, включаючи інформування постраждалих осіб та відповідних органів, а також вжиття заходів для мінімізації шкоди та запобігання подібним інцидентам у майбутньому.

Один із ключових викликів полягає в тому, щоб знайти баланс між забезпеченням конфіденційності та ефективністю роботи ШІ-систем. Занадто суворі заходи безпеки можуть ускладнювати або сповільнювати роботу систем, тоді як недостатній захист може призвести до витоку даних або інших серйозних наслідків.

З розвитком технологій з'являються нові загрози, пов'язані з обробкою персональних даних. Державні органи повинні бути готові адаптувати свої системи захисту до нових викликів, включаючи загрози, що виникають з використанням нових технологій, таких як квантові обчислення.

Оскільки проблеми конфіденційності та захисту персональних даних часто виходять за межі однієї країни, важливо співпрацювати на міжнародному рівні для розробки та впровадження єдиних стандартів і практик захисту.

Тому, ми робимо висновок, що конфіденційність та захист персональних даних є критичними аспектами використання ШІ в публічному управлінні. Вони вимагають комплексного підходу, що включає застосування сучасних технологій, розробку правових регуляцій та забезпечення прозорості для громадян. Лише за умов дотримання високих стандартів конфіденційності можна забезпечити довіру громадян до державних інституцій, які використовують ШІ у своїй роботі.

Використовуючи можливості та засоби інформаційних технологій та ШІ в публічному управлінні, маємо уникати дискримінації та упередженості стосовно громадян. Алгоритми ШІ можуть містити упередженість, якщо вони навчені

на основі історичних даних, що відображають існуючі соціальні нерівності. Це може призводити до дискримінації певних груп населення. Важливо розробляти такі алгоритми з урахуванням принципу справедливості та соціальної рівності.

Штучний інтелект (ШІ) має потенціал значно покращити ефективність публічного управління, однак його використання також несе ризик упередженості та дискримінації. Якщо алгоритми ШІ не розроблені або не впроваджені належним чином, вони можуть відтворювати або навіть посилювати існуючі соціальні нерівності. Тому важливо розробляти та використовувати такі системи, які враховують етичні принципи та мінімізують ризики дискримінації.

Розглянемо основні джерела дискримінації та упередження в ШІ.

1. Якість та репрезентативність даних Алгоритми ШІ навчаються на основі даних, які можуть містити упередження, відображаючи історичну дискримінацію або нерівність. Наприклад, якщо дані про минулі рішення стосовно працевлаштування містять упередження щодо певної етнічної групи, ШІ може навчитися повторювати ці упередження.

2. Алгоритмічне упередження Алгоритми, розроблені без належного врахування етичних аспектів, можуть бути упередженими щодо певних груп населення. Це може бути пов'язано з неправильною інтерпретацією даних або з тим, що алгоритми не були перевірені на відповідність етичним стандартам.

3. Системні чинники Системи ШІ можуть відтворювати упередження, що існують у суспільстві або у самих організаціях, які їх впроваджують. Якщо органи публічного управління мають вкорінені упередження або дискримінаційні практики, ШІ може посилити ці тенденції.

Умови уникнення дискримінації та упередження

1. Репрезентативність та баланс даних. Для уникнення дискримінації важливо використовувати збалансовані та репрезентативні дані. Це означає, що дані мають відображати різноманітність населення та включати інформацію про всі соціальні групи. Важливо також уникати використання даних, які можуть прямо або опосередковано викликати дискримінацію, наприклад, дані про расу, гендер або соціально-економічний статус, якщо це не є необхідним для конкретної задачі.

2. Регулярний аудит та тестування. Алгоритми ШІ мають підлягати регулярному аудиту та тестуванню для виявлення потенційних упереджень. Це включає проведення аналізу на різних етапах розробки та впровадження алгоритму, зокрема перевірку на відповідність етичним стандартам та відсутність дискримінаційних практик.

3. Прозорість алгоритмів. Прозорість є ключовим аспектом у боротьбі з упередженнями. Громадяни та незалежні експерти повинні мати доступ до інформації про те, як працює алгоритм, які дані він використовує, і як приймаються рішення. Це дозволяє вчасно виявляти та виправляти потенційні джерела дискримінації.

4. Етичні рамки та стандарти. Уряди та органи публічного управління мають впроваджувати чіткі етичні рамки для розробки та використання ШІ. Це включає створення національних або міжнародних стандартів, які регулюють використання алгоритмів, запобігають упередженням та забезпечують рівноправність усіх громадян.

5. Залучення різноманітних експертів. Важливо залучати до розробки та впровадження ШІ-систем фахівців з різних галузей, включаючи експертів із соціальних наук, правознавців та представників громадських організацій. Це допомагає забезпечити, що система враховує різні аспекти соціальної справедливості та права людини.

6. Освітні програми та підвищення обізнаності. Державні службовці та розробники ШІ повинні бути навчені розпізнавати та уникати упереджень. Освітні програми, що включають етичні аспекти використання ШІ, можуть значно знизити ризик впровадження дискримінаційних алгоритмів.

На нашу думку, уникнення дискримінації та упередження при використанні ШІ в публічному управлінні вимагає комплексного підходу, що включає правильний вибір і обробку даних, розробку прозорих і етичних алгоритмів, регулярний аудит та навчання. Забезпечення цих умов є необхідним для збереження довіри громадян і створення справедливих та інклюзивних систем публічного управління.

Одним із найважливіших аспектів використання ШІ в публічному управлінні є питання відповідальності. Хто нестиме відповідальність за рішення, прийняті на основі ШІ? Це питання потребує чіткого регулювання на законодавчому рівні.

Використання штучного інтелекту (ШІ) у публічному управлінні приносить значні переваги, зокрема підвищення ефективності та точності прийняття рішень. Однак, впровадження таких технологій також породжує питання щодо відповідальності за рішення, які приймаються ШІ. Хто повинен нести відповідальність у разі,

якщо рішення ШІ виявляться помилковими або призведуть до негативних наслідків? Відповідь на це питання є критично важливою для забезпечення довіри громадян до систем публічного управління, які використовують ШІ.

Розробники алгоритмів та постачальники ШІ-систем відіграють ключову роль у забезпеченні точності та етичності роботи ШІ. Вони несуть відповідальність за якість коду, що використовується, належну підготовку алгоритмів, а також за тестування та перевірку систем перед їх впровадженням. У разі виявлення серйозних недоліків або упереджень у роботі ШІ, відповідальність може лежати на розробниках, які не забезпечили належної якості продукту.

Органи публічного управління, які використовують ШІ, несуть кінцеву відповідальність за результати його роботи. Вони мають бути готовими відповідати за наслідки рішень, прийнятих ШІ, особливо якщо ці рішення впливають на життя громадян. Це означає, що державні установи повинні мати чіткі процедури для перевірки, моніторингу та, за потреби, коригування рішень ШІ.

У деяких випадках конкретні посадовці можуть бути відповідальними за роботу ШІ-системи або за нагляд за її використанням. Це можуть бути ІТ-спеціалісти, юристи або керівники відділів, відповідальних за впровадження технологій. Їхня відповідальність полягає в забезпеченні того, щоб ШІ працював відповідно до етичних норм і законодавства.

Механізми забезпечення відповідальності

1. Правові рамки Необхідно створити та підтримувати правову базу, яка визначає, хто несе відповідальність за рішення ШІ. Це може включати законодавство, що регулює впровадження та використання ШІ у публічному секторі, визначення стандартів для відповідності та можливі санкції у разі порушень.

2. Прозорість та підзвітність Для забезпечення відповідальності важливо впроваджувати прозорі процеси прийняття рішень та механізми підзвітності. Це включає звітування перед громадськістю, оприлюднення методологій та результатів використання ШІ, а також надання можливості оскарження рішень, прийнятих за допомогою ШІ.

3. Аудит та незалежний нагляд Регулярний аудит роботи ШІ-систем, а також незалежний нагляд з боку зовнішніх організацій чи експертів, є важливими для виявлення потенційних проблем на ранніх стадіях. Ці механізми можуть включати перевірку якості даних, з яких навчається ШІ, відповідності прийнятих рішень законодавству та етичним нормам, а також аналіз можливих наслідків для громадян.

4. Впровадження механізмів корекції та компенсації У разі, якщо ШІ приймає помилкове рішення, державні органи мають бути готові оперативно виправити ситуацію та, за необхідності, надати компенсацію постраждалим громадянам. Це вимагає створення механізмів швидкого реагування та вирішення конфліктів.

5. Етичні комітети та ради Органи публічного управління можуть створювати етичні комітети або ради, які будуть відповідати за розгляд складних питань, пов'язаних з використанням ШІ, включаючи питання відповідальності. Такі комітети можуть включати експертів із різних галузей, включаючи права людини, технології та соціальні науки.

Відповідальність за рішення, прийняті ШІ у публічному управлінні, повинна бути чітко визначена та розподілена між розробниками, постачальниками, державними установами та окремими посадовцями. Забезпечення належних механізмів правового регулювання, прозорості, аудиту та етичного нагляду є необхідним для запобігання негативним наслідкам та збереження довіри громадян до публічних послуг. В умовах швидкого розвитку технологій, державні установи повинні бути готові адаптувати свої практики та підходи до відповідальності, щоб відповідати викликам сучасного світу.

Говорячи про безпекові аспекти використання ШІ, перш за все ми маємо на увазі процес захисту від кібератак. Системи ШІ, як і будь-які інші інформаційні системи, можуть стати об'єктом кібератак, що може призвести до компрометації даних або підриву довіри до публічних інституцій. Необхідно впроваджувати найсучасніші засоби кіберзахисту для мінімізації цих ризиків. Штучний інтелект (ШІ) активно впроваджується в публічне управління з метою підвищення ефективності, швидкості та точності прийняття рішень. Проте разом із цими перевагами використання ШІ в публічному секторі також породжує нові виклики в сфері кібербезпеки. Захист від кібератак є надзвичайно важливим аспектом, оскільки атаки на ШІ-системи можуть призвести до серйозних наслідків, включаючи витік конфіденційних даних, маніпуляцію рішеннями або навіть паралич критично важливих державних функцій.

Основні кіберзагрози, пов'язані з використанням ШІ

1. Атаки на цілісність даних. ШІ-системи залежать від даних, на яких вони навчаються та з якими працюють. Атаки на цілісність даних

можуть призвести до того, що ШІ прийматиме хибні рішення на основі підроблених або маніпульованих даних. Наприклад, змінені дані можуть спотворити прогнозування або оцінку ризиків, що, у свою чергу, може спричинити небажані наслідки для громадян та держави.

2. Атаки на моделі машинного навчання. Хакери можуть намагатися зламати або маніпулювати моделями машинного навчання, використовуючи методи, такі як «атаки на вразливі місця» (adversarial attacks), коли вхідні дані спеціально модифікуються, щоб викликати помилкову відповідь алгоритму. Це може використовуватися для обходу систем безпеки або маніпуляції з результатами роботи ШІ.

3. Неавторизований доступ до ШІ-систем. ШІ-системи можуть стати мішенню для кібератак з метою отримання несанкціонованого доступу до конфіденційної інформації або для керування системою. Це може включати крадіжку особистих даних громадян, доступ до урядових документів або використання ШІ для шкідливих цілей.

4. Зловживання автономними системами. Деякі ШІ-системи в публічному управлінні можуть діяти автономно, без постійного нагляду людини. Якщо ці системи будуть скомпрометовані, вони можуть виконувати небажані дії, що може призвести до серйозних наслідків, таких як порушення громадського порядку або навіть шкода національній безпеці.

Стратегії захисту від кібератак при використанні ШІ

• Забезпечення цілісності даних. Одним із ключових елементів захисту є забезпечення цілісності та достовірності даних, які використовуються для навчання та роботи ШІ-систем. Це може включати застосування шифрування даних, використання криптографічних хеш-функцій для перевірки цілісності даних, а також впровадження багатофакторної аутентифікації для доступу до чутливих даних.

• Захист моделей машинного навчання. Для захисту моделей машинного навчання від атак слід використовувати методи захисту, такі як регулярне оновлення та перевірка моделей, впровадження алгоритмів, стійких до атак на вразливі місця, та проведення тестування на різні типи загроз. Важливо також впроваджувати механізми виявлення аномалій у вхідних даних, щоб вчасно виявляти можливі атаки.

• Контроль доступу та авторизація. Важливо впроваджувати суворі заходи контролю доступу до ШІ-систем, включаючи використання багатофакторної автентифікації, обмеження доступу до системи лише для уповноважених осіб, а також застосування систем моніторингу дій користувачів. Це дозволить виявляти та запобігати несанкціонованому доступу до систем.

• Аудит та регулярне оновлення. Проведення регулярних аудитів ШІ-систем на предмет вразливостей, а також постійне оновлення програмного забезпечення є критично важливими для забезпечення безпеки. Це включає впровадження сучасних засобів захисту від кіберзагроз, таких як системи виявлення вторгнень (IDS), антивірусне програмне забезпечення та захист від DDoS-атак.

• Розробка планів реагування на інциденти. Органи публічного управління повинні мати чітко розроблені плани реагування на кіберінциденти, включаючи протоколи для оперативного виявлення, ізоляції та нейтралізації атак. Це також включає планування відновлення після атак, що дозволяє швидко повернутися до нормальної роботи.

• Навчання та підвищення обізнаності. Освіта та підвищення обізнаності співробітників щодо кіберзагроз і методів захисту є важливими елементами кібербезпеки. Державні службовці та оператори ШІ-систем повинні регулярно проходити навчання з безпеки, щоб знати про новітні загрози та методи їхнього запобігання.

• Міжнародне співробітництво. Оскільки кіберзагрози мають глобальний характер, важливою є співпраця з міжнародними організаціями та іншими державами для обміну інформацією про загрози, кращі практики та координацію зусиль у боротьбі з кібератаками.

• Захист від кібератак при використанні ШІ в публічному управлінні є надзвичайно важливим для забезпечення стабільності та безпеки державних функцій. Це вимагає комплексного підходу, включаючи захист даних та моделей, суворий контроль доступу, регулярний аудит та оновлення систем, а також навчання персоналу. Забезпечення кібербезпеки не тільки захищає державні інституції від загроз, але й сприяє збереженню довіри громадян до уряду, що є основою ефективного публічного управління.

Використання ШІ для моделювання ризиків і прогнозування можливих загроз може значно підвищити безпеку публічного управління. Такі системи можуть бути корисними для запобігання кризам і забезпечення надійної роботи державних служб.

Штучний інтелект (ШІ) є потужним інструментом, який може значно підвищити ефективність публічного управління через автоматизацію

процесів, аналіз великих обсягів даних та покращення прогнозування. Однак, разом із цими можливостями, використання ШІ також несе певні ризики, які необхідно враховувати та контролювати. Моделювання ризиків та прогнозування при використанні ШІ у публічному управлінні є важливими для мінімізації потенційних негативних наслідків та підвищення надійності рішень.

Основні види ризиків, пов'язані з використанням ШІ

1. Технічні ризики Включають можливі помилки або несправності в роботі алгоритмів ШІ, що можуть призвести до неправильних рішень або невідповідних дій. Наприклад, недосконаліть алгоритмів машинного навчання може призводити до упередженості або неправильного трактування даних.

2. Етичні ризики Використання ШІ може викликати етичні дилеми, наприклад, щодо конфіденційності даних, справедливості рішень або дотримання прав людини. Також існує ризик створення систем, які можуть підривати суспільну довіру або сприяти дискримінації певних груп населення.

3. Юридичні ризики Порушення законодавства через використання ШІ, особливо у сфері захисту персональних даних, може призвести до юридичних наслідків для державних установ. Це включає ризик судових позовів або адміністративних санкцій.

4. Ризики кібербезпеки Як уже зазначалося, ШІ-системи можуть стати об'єктом кібератак, що може призвести до витоку конфіденційної інформації або маніпуляцій із даними, на основі яких приймаються рішення.

5. Соціальні ризики Використання ШІ може мати негативний вплив на ринок праці, призводити до зменшення робочих місць або змінювати соціальні відносини у суспільстві, що потребує уваги з боку уряду.

Моделювання ризиків є процесом ідентифікації, оцінки та аналізу потенційних загроз, пов'язаних з використанням ШІ, з метою розробки стратегій для їхнього зниження або запобігання.

На першому етапі необхідно визначити всі можливі ризики, пов'язані з використанням конкретних ШІ-систем у публічному управлінні. Це може включати як внутрішні ризики (наприклад, технічні несправності), так і зовнішні (наприклад, загрози кібербезпеці або соціальні наслідки).

Після ідентифікації ризиків їх слід оцінити за критеріями ймовірності та можливих наслідків. Ця оцінка допомагає визначити пріоритети у вирішенні ризиків та вибрати відповідні заходи для їхнього контролю. Наприклад, оцінка ризику помилкового рішення ШІ може включати аналіз того, наскільки велика ймовірність помилки та які будуть її наслідки для громадян або державних інституцій. Важливо враховувати, що різні ризики можуть бути взаємозалежними. Наприклад, технічні несправності можуть спричинити проблеми з кібербезпекою, а етичні ризики можуть викликати юридичні наслідки. Моделювання таких взаємозв'язків допомагає створити більш комплексний підхід до управління ризиками.

На основі результатів оцінки ризиків необхідно розробити стратегії для їхньої мінімізації або уникнення. Це може включати технічні заходи (наприклад, покращення алгоритмів або посилення кібербезпеки), організаційні зміни (наприклад, створення спеціалізованих відділів для управління ризиками) або правові заходи (наприклад, впровадження нових нормативних актів).

Прогнозування використання ШІ включає аналіз тенденцій та можливих сценаріїв розвитку технологій у майбутньому для прийняття стратегічних рішень у публічному управлінні.

Важливо аналізувати поточні тенденції у розвитку ШІ, зокрема в таких сферах, як машинне навчання, великі дані, автоматизація процесів. Це дозволяє прогнозувати, як ці технології будуть розвиватися і яким чином вони можуть бути використані в публічному управлінні.

Прогнозування включає аналіз того, як впровадження ШІ вплине на суспільство, ринок праці, економіку та соціальні відносини. Наприклад, важливо оцінити, які професії можуть бути замінені автоматизацією та які заходи слід вжити для підтримки працівників, що втратили роботу. Окрім короткострокових ризиків, слід також враховувати довгострокові перспективи, такі як можливість значних соціальних змін або вплив на міжнародні відносини. Це дозволяє приймати рішення, які будуть актуальними не тільки сьогодні, але й у майбутньому.

Використання сценарного підходу для прогнозування майбутнього дозволяє розробити різні можливі сценарії розвитку ШІ та їхнього впливу на публічне управління. Це допомагає підготуватися до різних варіантів розвитку подій і розробити гнучкі стратегії реагування.

Моделювання ризиків та прогнозування є невід'ємними елементами управління впровадженням ШІ в публічному управлінні. Вони дозволяють ідентифікувати та мінімізувати потенційні загрози, а також прогнозувати майбутні

можливості та виклики. Комплексний підхід до оцінки ризиків та прогнозування дозволяє забезпечити стійкість, ефективність та етичність використання ШІ у державному секторі, що є ключовим для довгострокового успіху цих технологій.

Безпекові загрози, пов'язані з ШІ, часто мають глобальний характер, що вимагає міжнародного співробітництва. Спільна розробка стандартів і протоколів безпеки може забезпечити більш ефективний захист від потенційних загроз.

Використання штучного інтелекту (ШІ) у публічному управлінні стрімко поширюється по всьому світу, що створює нові можливості, але й водночас породжує численні виклики, зокрема у сфері безпеки. Ці виклики мають глобальний характер, що робить міжнародне співробітництво критично важливим для забезпечення безпеки ШІ на національному та міжнародному рівнях. У цьому контексті співпраця між країнами, міжнародними організаціями та іншими зацікавленими сторонами стає необхідною для розробки та впровадження ефективних стандартів і практик безпеки.

Основні напрями міжнародного співробітництва

1. Розробка міжнародних стандартів і нормативних актів Одна з ключових форм міжнародного співробітництва у сфері безпеки ШІ полягає у спільній розробці міжнародних стандартів та нормативних актів, які регулюють використання ШІ в публічному управлінні. Ці стандарти можуть включати технічні вимоги до безпеки, етичні норми, принципи захисту даних та управління ризиками. Такі стандарти допомагають забезпечити узгоджений підхід до використання ШІ в різних країнах та полегшують співпрацю між державами.

2. Обмін інформацією про загрози та найкращі практики Міжнародне співробітництво включає обмін інформацією про кіберзагрози, вразливості ШІ-систем та найкращі практики щодо їхнього запобігання. Це дозволяє країнам швидше реагувати на нові загрози, а також впроваджувати перевірені підходи до забезпечення безпеки. Обмін інформацією може здійснюватися через міжнародні організації, такі як ООН, ЄС, НАТО, а також через двосторонні або багатосторонні угоди.

3. Спільні дослідницькі проєкти та ініціативи Країни можуть співпрацювати у проведенні спільних дослідницьких проектів, спрямованих на вивчення безпеки ШІ та розробку нових рішень для підвищення захисту. Такі ініціативи можуть включати розробку нових методів кібербезпеки, аналіз етичних викликів використання ШІ, а також створення інноваційних підходів до управління ризиками. Спільні дослідження сприяють підвищенню рівня знань та розвитку нових технологій у сфері безпеки ШІ.

4. Міжнародні навчальні програми та підготовка кадрів Важливим аспектом міжнародного співробітництва є навчання та підготовка кадрів у сфері безпеки ШІ. Країни можуть обмінюватися досвідом у навчанні фахівців, організовувати спільні тренінги та семінари, а також створювати міжнародні навчальні програми. Це сприяє підвищенню кваліфікації працівників державного сектору та дозволяє краще підготуватися до викликів, пов'язаних з використанням ШІ.

5. Створення міжнародних платформ для співпраці Для ефективної координації зусиль у сфері безпеки ШІ важливо створювати міжнародні платформи для обговорення та співпраці. Це можуть бути спеціалізовані форуми, конференції, робочі групи або консорціуми, які об'єднують представників різних країн, міжнародних організацій, наукових установ та приватного сектору. Такі платформи дозволяють обговорювати актуальні питання, розробляти спільні рішення та зміцнювати міжнародні зв'язки.

Висновки. Використання штучного інтелекту в публічному управлінні має величезний потенціал для покращення державних послуг і підвищення ефективності управлінських процесів. Однак для досягнення цих цілей необхідно забезпечити дотримання етичних стандартів та безпеки. Це вимагає комплексного підходу, що включає розробку правових рамок, впровадження технологічних засобів захисту та міжнародного співробітництва. Міжнародне співробітництво у сфері безпеки ШІ в публічному управлінні є важливим інструментом для забезпечення глобальної безпеки та ефективності використання ШІ. Спільна розробка стандартів, обмін інформацією, співпраця у дослідженнях та навчанні, а також створення міжнародних платформ для діалогу допомагають країнам краще підготуватися до викликів, пов'язаних з використанням ШІ. Однак для досягнення успіху необхідно долати виклики, пов'язані з різницею в законодавстві, технологічним розвитком та геополітичними розбіжностями, зміцнюючи довіру та розвиваючи міжнародне партнерство.

Список літератури:

1. O. Obolenskyi (2023). Course of lectures on the subject Conceptual principles of public management and administration.
2. Кравчук О.Ю. Штучний інтелект у фокусі стратегічного планування публічного управління. *Суспільство і національні інтереси*. № 4(4), 2024.

Kravchuk O.Yu. ENSURING ETHICAL STANDARDS AND SAFETY OF USING ARTIFICIAL INTELLIGENCE IN PUBLIC ADMINISTRATION

The article examines key aspects of ensuring ethical standards and security in the use of AI in public administration, including issues of confidentiality, transparency, accountability and avoidance of discrimination. The authors also analyze international experience and provide recommendations for the implementation of best practices in this area.With the development of artificial intelligence (AI) technologies and its integration into various spheres of public life, including public administration, new ethical and security challenges arise. In the modern world, where artificial intelligence (AI) is becoming an integral part of social life, its implementation in the field of public administration is of great interest. However, along with the potential benefits that AI can bring to improve the effectiveness of management, there are a number of challenges related to the ethics and security of its use. This article is devoted to the analysis of the main ethical standards and security aspects that must be taken into account during the integration of AI into public administration. In addition, the article considers the problems of responsibility and transparency in decision-making based on AI algorithms, which can affect the rights and freedoms of citizens. The article also analyzes international practices and recommendations for ensuring the ethical use of AI, including the development of relevant legal acts and the creation of special regulatory bodies.

The authors emphasize the need to develop ethical codes and standards for public administration taking into account the specifics of AI application, as well as the importance of raising awareness among public officials and the general public about the potential risks associated with AI. Based on the analysis, the article offers recommendations for the implementation of comprehensive approaches to ensuring ethical standards and security of AI, which include technical, organizational, and regulatory measures. The conclusions of the article are aimed at promoting the formation of a responsible approach to the use of AI in public administration, which will maximize its positive impact and minimize possible negative consequences.

Key words: *artificial intelligence, public administration, ethics, security, confidentiality, transparency, public administration.*