DOI 10.31891/2307-5732-2024-337-3-55
УДК 004.056: 004.9

ЛОЗОВСЬКИЙ РОСТИСЛАВ
https://orcid.org/0009-0004-9611-9424
e-mail: rawrshah@gmail.com
МОРОЗ АНТОН
https://orcid.org/0009-0001-1942-8432
e-mail: airmoroz26@gmail.com

ВПЛИВ ШТУЧНОГО ІНТЕЛЕКТУ НА СТРАТЕГІЇ ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД НОВИХ ТИПІВ КІБЕРЗАГРОЗ

У статті визначено роль і значення штучного інтелекту в забезпеченні кібербезпеки, а також детально розглянуто методи його застосування. Окреслено переваги та ключові пріоритети впровадження штучного інтелекту в цю сферу. Проведено аналіз недоліків і ризиків, пов'язаних із використанням технологій штучного інтелекту хакерами в кібербезпеці. Вивчено інновації, досягнуті в практичному впровадженні генеративного штучного інтелекту, зокрема ChatGPT (Generative Pre-trained Transformer). Означено правові основи регулювання штучного інтелекту в галузі кібербезпеки в Україні. Виокремлено загрозливі тенденції у використанні технологій штучного інтелекту на основі звіту Європолу за 2023 рік. Огляд законодавчих ініціатив ЄС, що стосуються регулювання штучного інтелекту, зокрема в контексті кібербезпеки, також включено в дослідження. Узагальнено рекомендації щодо вдосконалення правових норм для ефективного використання технологій штучного інтелекту в сфері кібербезпеки, особливо в умовах воєнного стану в Україні.

Ключові слова: штучний інтелект, кібербезпека, кібератака, кіберзагроза, національна безпека, інформаційні технології, машинне навчання.

LOZOVSKYI ROSTYSLAV
MOROZ ANTON

IMPACT OF ARTIFICIAL INTELLIGENCE ON INFORMATION SYSTEMS PROTECTION STRATEGIES AGAINST NEW TYPES OF CYBER THREATS

This article explores the impact of artificial intelligence (AI) on strategies for protecting information systems against new types of cyber threats, with a focus on current methods and technologies. It examines not only the key advantages of using AI, such as improved threat detection accuracy, faster response to cyberattacks, and automation of security management processes, but also the drawbacks and challenges associated with its application. Innovative aspects of generative artificial intelligence, including ChatGPT, and its impact on the development of new attack and defense methods are studied. Legal aspects of regulating AI in the field of cybersecurity in Ukraine are investigated, along with the primary threats arising from the implementation of such technologies, based on the 2023 Europol report. The review of EU legislative initiatives concerning AI regulation, particularly in the context of cybersecurity, provides essential insights. The article also offers recommendations for improving legal frameworks to better utilize AI in cybersecurity, especially under the conditions of martial law in Ukraine, aiming to enhance national security and protect information technologies.

In addition to examining the benefits and limitations of AI in cybersecurity, this article delves into the practical challenges organizations face when integrating AI solutions. It addresses issues such as the need for substantial computational resources, the potential for algorithmic biases, and the complexities of ensuring data privacy and ethical use of AI technologies. The discussion highlights the importance of balancing technological advancement with robust ethical standards and regulatory oversight to mitigate risks associated with the misuse of AI.

Furthermore, the article explores the potential future developments in AI that could influence cybersecurity strategies. It considers emerging trends such as the evolution of machine learning algorithms, the integration of AI with other advanced technologies like quantum computing, and the role of international collaboration in addressing global cybersecurity threats. By analyzing these aspects, the article aims to provide a comprehensive understanding of how AI can be leveraged to enhance cybersecurity while also addressing the associated challenges and risks.

Keywords: artificial intelligence, cybersecurity, cyberattack, cyber threat, national security, information technologies, machine learning.

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

У сучасному цифровому середовищі кіберзагрози постійно еволюціонують, стаючи дедалі складнішими та агресивнішими. Інформаційні системи, які є критично важливими для функціонування сучасних організацій, стикаються з новими викликами, що вимагають вдосконалення методів захисту. Впровадження штучного інтелекту (ШІ) в область кібербезпеки стає однією з найзначніших тенденцій, що обіцяє революціонізувати стратегії захисту від кіберзагроз.

Штучний інтелект, завдяки своїм можливостям обробки великих обсягів даних та швидкому навчанні на основі нових загроз, має потенціал для значного покращення процесів виявлення і нейтралізації кіберзлочинів. Однак разом із перевагами, які приносить ШІ, виникають і нові виклики, такі як етичні питання, ризики зловживання технологією та необхідність адаптації існуючих стратегій кібербезпеки.

Встановлено, яким чином штучний інтелект впливає на сучасні стратегії захисту інформаційних систем від нових типів кіберзагроз. Вона розгляне ефективність застосування ШІ у виявленні та нейтралізації кіберзагроз, проаналізує переваги та обмеження інтеграції ШІ у системи кіберзахисту, а також оцінить потенційні ризики та етичні аспекти, пов'язані з використанням штучного інтелекту в цій критично важливій сфері. Через це дослідження буде можливим краще зрозуміти, як сучасні технології ШІ можуть вплинути на подальший розвиток стратегій кібербезпеки, забезпечуючи таким чином ефективний захист від нових і дедалі більш небезпечних кіберзагроз.

Аналіз досліджень та публікацій

Необхідно зазначити, що деякі теоретичні та практичні аспекти використання штучного інтелекту (ШІ) в кібернетичній сфері як в Україні, так і за кордоном вже були розглянуті в науковій літературі. Наприклад, технічні аспекти та особливості застосування систем ШІ для забезпечення кібербезпеки досліджували О. Неретін і В. Харченко [1], В. Савченко і О. Шаповаленко [2], а також І. Стьопочкіна і О. Новіков [3]. Перспективні можливості ШІ як важливого інструменту для автоматизованої і оперативної реакції на нові та модифіковані кіберзагрози аналізував С. Шаров [4]. С. Цяпа [5] провів огляд закордонних законодавчих ініціатив щодо стратегічного використання технологій ШІ в кібербезпеці. У закордонній науковій літературі роль і значення ШІ в кібербезпеці, а також можливі напрями його подальшого розвитку досліджували Т. Сіпола [6], Р. Мостіну [7], Р. Дас і Р. Сандхейн [8].

Однак питання застосування ШІ в сфері кібербезпеки досі недостатньо вивчені на науковому рівні. Це особливо помітно в умовах появи феномену генеративного ШІ, такого як ChatGPT у листопаді 2022 року, та тривалого (протягом останніх 18 місяців) правового режиму воєнного стану, що робить цю тематику ще більш актуальною для наукового дослідження.

Формулювання цілей статті

Метою роботи є: дослідити вплив штучного інтелекту на сучасні стратегії захисту інформаційних систем, зокрема шляхом аналізу ефективності ШІ у виявленні, нейтралізації та управлінні новими типами кіберзагроз, оцінити переваги та виклики інтеграції ШІ у кібербезпеку, розробити рекомендації щодо оптимізації стратегій захисту на основі сучасних технологій ШІ, а також визначити потенційні ризики та етичні питання, що виникають при використанні ШІ в цій сфері.

Виклад основного матеріалу

Працювати в кіберпросторі та відстежувати постійно зростаючі кіберзагрози вимагає залучення великої кількості висококваліфікованих спеціалістів. Водночас штучний інтелект може частково взяти на себе ці завдання, адже він значно швидше виявляє уразливості та генерує коди й алгоритми. Виявлення загроз і уразливостей набуло значних масштабів і стає проблемою для засобів захисту периметру та даних ІКТ-систем, що керуються людиною. Кібератаки стають набагато складнішими і дуже руйнівними, а також зростає ризик потрапляння небезпечних технологій до рук зловмисників та конкурентів.

У зв’язку зі стрімким розвитком Інтернету речей (IoT), хмарних обчислень, центрів обробки даних (ЦОД) та відповідно технологій обробки великих даних (big data), змінилася парадигма інформаційної безпеки (ІБ) – від захисту периметру корпоративної мережі, «хмари» або ЦОДу до захисту самих даних. Традиційно для безпечного доступу до ІТ-ресурсів або об'єктів використовуються віртуальні приватні мережі (VPN) як тунелі на основі криптографічних методів захисту інформації. Однак використання лише VPN несе ризики для безпеки. Проблема полягає в тому, що VPN базується на підході до безпеки через захист периметру. Користувачі підключаються через VPN-клієнт, але потрапляючи всередину периметру, часто отримують широкий доступ до інформаційних ресурсів і керування об'єктом. Кожного разу, коли пристрій або користувач автоматично отримують таку довіру, це створює загрозу для даних, додатків та інтелектуальної власності організації.

Окрім проблем, пов'язаних з використанням VPN для віддаленого доступу, мережеві оператори шукають оптимальний спосіб захисту додатків. Ситуація ускладнюється тим, що частина додатків розміщена в хмарі, а частина – локально, що ускладнює застосування єдиного методу контролю, особливо коли одні користувачі працюють в офісі, а інші – віддалено. Розгортання додатків у хмарі створює можливість для сканування з боку небажаних суб’єктів, що значно підвищує ризик. Виходом із цієї ситуації є використання нової технології, яка виходить за межі VPN – доступ з нульовою довірою (ZTNA, англ. Zero Trust Network Access), що з використанням штучного інтелекту та машинного навчання пропонує більш ефективне рішення для віддаленого доступу і також вирішує проблеми, пов'язані з доступом до додатків. Термін «нульова довіра» слід розуміти буквально. Ця модель безпеки передбачає, що жоден користувач або пристрій не вважаються надійними, і жодна транзакція не заслуговує на довіру без попередньої перевірки авторизації користувача та пристрою. Оскільки ZTNA виходить із того, що місце розташування не визначає рівень довіри, фізичне місцезнаходження користувача не має значення. Цей підхід застосовується незалежно від того, де фізично знаходиться користувач або пристрій. Оскільки будь-який пристрій потенційно може бути зараженим, а будь-який користувач здатний на шкідливу поведінку, політика доступу ZTNA відображає цю реальність. На відміну від традиційного VPN-тунелю з необмеженим доступом, ZTNA надає доступ до кожного додатка і робочого процесу лише після аутентифікації користувача та/або пристрою. Перш ніж отримати доступ, користувачі проходять верифікацію і аутентифікацію для доступу до додатку. Кожен пристрій також перевіряється при кожному доступі до додатку для забезпечення відповідності вимогам політики доступу.

Під час авторизації використовується різноманітна контекстна інформація, така як роль користувача, тип пристрою, відповідність пристрою вимогам, місцезнаходження, час і спосіб підключення пристрою або користувача до мережі чи ресурсу. Якщо використовується технологія ZTNA, то після введення користувачем необхідних облікових даних для багатофакторної аутентифікації та перевірки кінцевої точки, йому надається доступ з обмеженими правами. Користувач може отримати доступ лише до тих додатків, які потрібні йому для ефективного виконання роботи, і не більше. Контроль доступу не обмежується лише точкою входу. ZTNA

працює на основі ідентифікації, а не захисту певного сегмента мережі, що дозволяє політикам безпеки (ПБ) контролювати додатки та інші транзакції від початку до кінця. Завдяки високому рівню контролю доступу ZTNA є більш ефективним рішенням для кінцевих користувачів і забезпечує дотримання ПБ скрізь, де це необхідно. І хоча процес аутентифікації ZTNA забезпечує точки перевірки автентичності, на відміну від традиційної VPN, він не визначає, як саме повинна проходити ця аутентифікація. Зі впровадженням нових рішень для автентифікації їх можна легко інтегрувати в стратегію ZTNA. Нові рішення для перевірки автентичності можуть допомогти вирішити проблеми, пов'язані зі слабкими або вкраденими паролями та обліковими даними, підвищити безпеку деяких пристроїв Інтернету речей (IoT) або додати додаткові рівні перевірки для доступу до конфіденційної інформації або важливих ресурсів.

Забезпечення доступності об'єкта захисту є невід'ємною частиною забезпечення інформаційної безпеки, тому системи моніторингу продуктивності та доступності є обов'язковими інструментами при здійсненні моніторингу ІБ в ІТ-системах. Системи моніторингу продуктивності можуть використовуватися як окремо, так і бути джерелом подій для системи управління подіями – SIEM (англ. Security Information Event Management). Такі системи призначені для відстеження стану функціонування різноманітних мережевих сервісів та її вузлів (серверів, мережевого обладнання, додатків та інших), включаючи підсистеми ІБ, на основі різних критеріїв продуктивності та доступності. У таких рішеннях застосовуються такі основні методи контролю функціонування з точки зору забезпечення доступності систем:

- збір та агрегація різноманітних даних, показників та лічильників використання апаратних ресурсів системи, зазвичай через встановлених агентів на контрольованих вузлах або з використанням протоколу SNMP;
- аналіз та кореляція зібраних даних для визначення або попередження досягнення граничних значень показників продуктивності та доступності з метою реагування або запобігання нестандартним ситуаціям у функціонуванні систем;
- автоматизоване виконання заздалегідь запрограмованих тестів, що перевіряють функціонування різних параметрів сервісів за заданим сценарієм. Успішне виконання таких тестових сценаріїв дозволяє підтвердити доступність сервісів та систем на різних рівнях;
- автоматизоване реагування системи у вигляді виконання заданих скриптів, програм або задач при виявленні значних відхилень показників на етапі кореляції;
- генерування повідомлень про виявлені відхилення у продуктивності та доступності систем. Сповіщення може відображатися на екрані моніторингу інтерфейсу системи, а також направлятися через різні канали повідомлень – електронну пошту, GSM-шлюз, системи обміну миттєвими повідомленнями (наприклад, jabber) та інші;
- візуалізація зібраних даних у вигляді діаграм, що допомагають ідентифікувати аномалії або значні відхилення, відмінні від стандартної поведінки систем. Візуалізація включає також представлення даних у вигляді звітів;
- зберігання зібраних даних у базі даних.

Оскільки метою нашої статті є опис методів моніторингу інформаційної безпеки з використанням штучного інтелекту та машинного навчання, ми не будемо детально розглядати ці системи моніторингу з точки зору їх архітектури та особливостей функціонування.

Ще однією технологією забезпечення кібербезпеки, яка широко використовує функції штучного інтелекту та машинного навчання, є DLP. Ця абревіатура розшифровується як Data Loss Prevention (запобігання втраті даних) або Data Leakage Prevention (запобігання витоку даних). Найчастіше для продуктів цього класу використовують саме це скорочення, але зустрічаються й інші. Якщо ви стикаєтеся з абревіатурами ILP, ILDP, EPS або CMF, швидше за все, йдеться про систему безпеки, яка захищає від витоків даних. DLP-системи базуються на аналізі потоків даних, що перетинають периметр захищеної інформаційної системи. Якщо в цьому потоці виявляється конфіденційна інформація, активна компонента системи блокує передачу повідомлення (пакета, потоку, сесії). Основою "інтелектуальних функцій" DLP є технологія так званого контентного аналізу, яка включає:

Пошук за ключовими словами та словниками. Це передбачає пошук точних збігів текстових рядків, де можуть використовуватися спеціальні символи, що позначають групи символів. Не слід плутати пошук за ключовими словами з технологією лінгвістичного аналізу, яка також є в DLP і враховує різні словоформи. Офіцер безпеки може створювати власні словники під конкретні тематики або скористатися одним із попередньо встановлених словників.

Машинне навчання. Самонавчальна технологія DLP базується на алгоритмі Байєса та методі опорних векторів. У процесі аналізу різних документів технологія самостійно виділяє ознаки різних категорій конфіденційних даних. Чим більше конфіденційних документів система побачить на етапі навчання, тим вищою буде її результативність у щоденній роботі.

Цифрові відбитки Docu Prints. Технологія ґрунтується на порівнянні перехопленої інформації з зразками конфіденційних документів, зокрема оцифрованих документів за їх характерними деталями. Технологія особливо ефективна для контролю документів, обсяг яких значний, а вміст під час використання змінюється незначно. DLP є стійкою до модифікацій вихідних документів і максимально точно виявляє інформацію, яка повністю або частково збігається із заданими конфіденційними даними.

Аналіз графічних файлів (OCR). Технологія розпізнає конфіденційні дані, що містяться у скріншотах,

фотографіях, відсканованих, рукописних та інших графічних документах. Інтегровані в DLP OCR-модулі ABBYY FineReader і Google Tesseract витягують текст із зображень для подальшого аналізу та перевірки відповідності політикам безпеки DLP.

Графічні відбитки. Ця технологія використовується для виявлення таких елементів у графічних файлах, як підписи, печатки, бланки, а також зображення з певною структурою, наприклад, скани паспортів або водійських прав. Спотворення кольорів, нахил зображення, накладені елементи (наприклад, написи та фонова графіка на печатці) не заважають розпізнаванню образів. Окрім шаблонів і регулярних виразів, DLP виявляє передачу структурованої інформації – паспортних даних, адрес, номерів кредитних карток, банківських рахунків, URL-адрес, номерів телефонів та інших даних, а також будь-яких типів даних за заданими регулярними виразами.

Давайте порівняємо можливості двох останніх розглянутих нами систем – SIEM та DLP. Зазначимо, що одна система не замінює іншу, вони вирішують різні завдання. Простий приклад: співробітник кілька разів ввів неправильний пароль. SIEM не лише виявить ці дії, але й співставить фактори – скільки разів пароль введено неправильно? протягом якого часу? Система виявить загрозу інформаційній безпеці – хтось намагається підібрати пароль до облікових даних – і своєчасно сповістить про це. Проте без глибокого аналізу SIEM є мало корисною. DLP, у свою чергу, дозволяє деталізувати дані та з'ясувати подробиці інциденту. Такий симбіоз SIEM та DLP суттєво підвищує рівень інформаційного захисту організації та спрощує роботу служби безпеки.

Зараз розробники DLP активно інтегрують свої рішення з популярними SIEM-системами. Процес лише набирає обертів, і говорити про конкретні результати ще рано. Для користувачів інтеграція своїх даних у SIEM – це повільний процес, адже недостатньо просто передати дані сторонній системі, потрібно зрозуміти, які конкретні завдання здатна вирішувати така схема та як саме допоможе цей союз. Знадобиться ще не одна ітерація, щоб ця зв'язка почала вирішувати прикладні завдання – вибирати з величезного масиву даних необхідні й стати зручним і функціональним інструментом для вирішення реальних проблем, а не просто конструктором.

Проте, в чому основна проблема більшості SIEM-систем і DLP? У тому, що мало хто з корпоративних «офіцерів безпеки» розуміє, як використовувати величезний масив зібраної інформації (Big Data) для налаштування цих інтелектуальних систем. Робота з системою вимагає спеціальних знань і навичок (Data Science), і якщо їх немає, то SIEM і DLP перетворюються на дорогий і практично безкорисний для бізнесу інструмент. Це обставина є одночасно і уразливістю, і загрозою для корпоративних і державних систем (особливо тих, що зберігають і обробляють критично важливу інформацію), оскільки доводиться передавати Big Data стороннім фахівцям із Data Science, які й здійснюють машинне навчання цих інтелектуальних систем. Тому завдання розробників – зробити так, щоб симбіозна система була максимально дружньою до машинного навчання і зрозумілою корпоративним спеціалістам з інформаційної безпеки. Іншою альтернативою є те, щоб ці фахівці з Data Science надавали зобов'язання про нерозголошення і непередачу отриманих для машинного навчання даних стороннім організаціям, подібно до того, як це відбувається з інформацією, що становить державну таємницю.

Нова світова гонка технологій у найближчому майбутньому призведе до впровадження найсучасніших інновацій у військову сферу. Цим будуть займатися всі провідні світові держави, оскільки будь-яке відставання від конкурентів збільшує вразливість, яку буде дуже складно компенсувати звичайними конвенційними видами озброєнь. Крім того, поява нових технологій може призвести до значних змін у стратегіях, плануванні та організації діяльності збройних сил. Запобігти використанню штучного інтелекту у військових цілях неможливо. Штучний інтелект значно розширить можливості збору та аналізу даних, що дозволить отримати певні переваги в швидкості та якості обробки інформації. У сфері військової розвідки з'являться нові можливості та різноманітні джерела інформації, але й можливостей приховати правду від противника також побільшає. Штучний інтелект може доповнити інформаційний простір великим обсягом штучно створених даних, віртуальною правдою, що з одного боку заплутає потенційних противників, а з іншого – може створити додаткові політичні ризики.

Зараз навіть ті технології, що вже створені у сфері машинного навчання та штучного інтелекту, мають значний потенціал для забезпечення національної безпеки. Існуюча технологія розпізнавання образів може забезпечити високу ступінь автоматизації під час аналізу супутникових знімків та даних радарів. Штучний інтелект здатен підвищити ефективність роботи радіолокаційних станцій системи попередження про ракетний напад та системи обробки інформації на радіооптичних комплексах розпізнавання. Крім того, нинішня мініатюризація супутників та збільшення їх кількості на орбітах вимагатимуть технологій для швидкого розпізнавання. Ще масштабніші завдання стоять перед комплексами обробки інформації загоризонтних радіолокаторів, які використовують принцип просторового іоносферного поширення радіохвиль довжиною понад 10 метрів або дифракційного поверхневого поширення коротших радіохвиль. Ці радари «бачать» всі рухомі об'єкти, включно з цивільною технікою, тому важливо швидко розпізнати серед усіх отриманих тисяч і навіть мільйонів образів саме військові об'єкти, а також незвичну поведінку на землі та в повітрі. Це величезні масиви інформації та образів, з обробкою яких люди не зможуть впоратися без допомоги машин. До того ж військові отримають так звану «бібліотеку цілей», що допоможе системам розпізнавання та наведення. Якщо для протидії ПЗРК з інфрачервоною головкою самонаведення достатньо з борту літака або вертольота відстрілювати хибні теплові цілі, а проти радіолокаційної ставити перешкоди, то системи зі

штучним інтелектом, навіть якщо він знаходиться не в ракеті, а в руках оператора, «бачать» літальний апарат повністю. Крім того, у загоризонтних РЛС існує проблема несумісності зі стандартною системою радіолокаційного розпізнавання «свій-чужий», тому під час аналізу повітряної обстановки допомога штучного інтелекту буде дуже доречною. Штучний інтелект також можна використовувати для протидії радарам противника, вивчаючи їх роботу та підбираючи методи придушення радіосигналу.

Можливості автономних систем наразі обмежені. Незважаючи на те, що такі системи є "квазіавтономними", люди все одно повинні залишатися в контурі управління і безпосередньо приймати рішення щодо застосування зброї. Проте звичайна людина, у порівнянні з можливостями сучасної військової техніки, є слабкою, крихкою і недосконалою істотою, а в ланцюзі прийняття бойових рішень – ще й найповільнішою ланкою. Штучний інтелект покликаний повністю виключити людину з процесу прийняття рішень, одночасно зберігаючи життя військовослужбовців.

У бойових умовах перевага буде за тими, хто прийме рішення швидше і завдасть удару першим, тому повністю автономні системи в майбутньому отримають значний розвиток. Ба більше, вже з'явилася концепція "контравтономності", згідно з якою штучний інтелект, зазнавши нападу, але залишившись непошкодженим, дуже швидко навчиться, зробить висновки і завдасть смертельного удару противнику. Можливості застосування тактичної зброї зі штучним інтелектом різноманітні. Це і безпілотні літальні апарати, і бронемашини, і ракетні катери, які самостійно знаходять цілі та приймають рішення щодо їх знищення. Наразі відбувається швидке зниження вартості безпілотників і дронів, а їхнє виробництво стає масовим. Використання штучного інтелекту допоможе об'єднувати тисячі дронів у величезний керований "рій", здатний до масової атаки.

Штучний інтелект також може бути інтегрований у технології державного управління та зміцнення влади, ставши інструментом внутрішньої політики. Він також стане помічником державних органів у керуванні екологічними ризиками та запобіганні техногенним катастрофам. Прогрес у створенні штучного інтелекту матиме потужний вплив на економіку і може призвести до нової промислової революції. Держава, яка першою впровадить такі технології, здобуде економічну, інформаційну, а можливо, і військово-політичну перевагу над іншими країнами.

Таблиця 1

Аналіз міфів про аспекти застосування штучного інтелекту

Міфи про ШІ	Реальність	Приклади та факти
ШІ — це просто розумні машини, які можуть замінити людину в будь-якій сфері.	Хоча ШІ може виконувати конкретні завдання, він не має універсальних когнітивних здібностей і не може повністю замінити людину у складних та багатогранних сферах, таких як творчість та стратегічне мислення.	ШІ може добре справлятися з аналізом даних і розпізнаванням образів (наприклад, у медичних діагнозах), але він не може замінити людське чуття та інтуїцію, які важливі в мистецтві або психології.
ШІ може стати свідомим і розвинути власні цілі.	На даний момент ШІ не має свідомості або власних цілей. Він працює на основі алгоритмів і даних, що надаються людьми.	Сучасні системи ШІ, такі як чат-боти або системи рекомендацій, не мають власних думок або бажань. Вони діють згідно з програмуванням і алгоритмами, розробленими людьми.
ШІ завжди об'єктивний і позбавлений упереджень.	ШІ може успадковувати упередження, що є в даних, на яких він навчався. Це може призвести до несправедливих або неточних рішень.	Розпізнавання обличчя на основі ШІ виявилося менш точним для представників різних расових груп через недостатність даних для деяких груп, що веде до упереджень у системах безпеки.
ШІ може контролювати світ і приймати глобальні рішення.	ШІ може допомагати в прийнятті рішень, але остаточні рішення приймаються людьми. ШІ не має автономного контролю і потребує людського нагляду і регулювання.	Автоматизовані торгові системи можуть здійснювати операції на фондовому ринку, але їхні рішення контролюються людьми, і є випадки, коли людський фактор все ще відіграє ключову роль.

Таблиця 1 містить популярні міфи про штучний інтелект і відповідні реальні факти, що спростовують ці міфи. Для кожного міфу наводяться конкретні приклади та факти, які допомагають зрозуміти, чому ці міфи не відповідають дійсності. Таблиця допомагає розкрити переваги та обмеження сучасних технологій штучного інтелекту, а також підкреслює важливість точного розуміння їх можливостей і обмежень.

В таблиці 2 проаналізовано основні загрози, пов'язані з використанням штучного інтелекту, та наводить конкретні приклади і факти, що ілюструють ці загрози в реальному світі. Таблиця охоплює питання приватності, безпеки, можливих негативних наслідків автоматизації, а також етичних і моральних проблем, що можуть виникнути внаслідок застосування ШІ. Мета таблиці — надати чітке уявлення про потенційні ризики та способи їх мінімізації в контексті технологій штучного інтелекту.

Таблиця 2

Загрози штучного інтелекту і їх реальні прояви

Загрози ШІ	Реальність	Приклади та факти
Загрози приватності та безпеки.	ШІ може бути використано для збору та аналізу великих обсягів особистих даних, що може загрожувати приватності. Неправильний захист даних або уразливості систем можуть призвести до витоків або зловживань.	Використання ШІ для трекінгу та аналізу даних користувачів у мобільних додатках може призвести до порушення приватності, якщо дані не захищені належним чином.
Автоматизація і втрата робочих місць.	Автоматизація може призвести до змін у структурі робочих місць і навіть їх скорочення в деяких галузях. Важливо розробити стратегії для перепідготовки працівників і створення нових можливостей.	Автоматизація виробничих процесів в автомобільній промисловості може призвести до скорочення робочих місць на конвеєрах, але також створює нові можливості в розробці і обслуговуванні автоматизованих систем.
Загрози від шкідливого використання ШІ.	ШІ може бути використано для кібератак, маніпулювання громадською думкою або розробки зброї. Необхідні заходи для регулювання та контролю за використанням ШІ, щоб зменшити ризики.	ШІ може бути використано для створення фальшивих новин або "deepfake" відео, які можуть маніпулювати громадською думкою або вчиняти шахрайство.
Етичні та моральні питання.	Питання етики та моралі в застосуванні ШІ включають проблеми прийняття рішень, пов'язані з правами людини та справедливістю. Важливо враховувати ці аспекти при розробці та впровадженні систем ШІ.	У сфері кримінального правосуддя алгоритми ШІ можуть бути використані для прогнозування рецидиву злочинців, але існують питання про те, наскільки ці системи є справедливими і прозорими.

Висновки з даного дослідження
і перспективи подальших розвідок у даному напрямі

Проблемам інформаційної безпеки та захисту суспільства від негативного впливу штучного інтелекту (ШІ) і машинного навчання (МО) приділяється значна увага в ряді досліджень. Основними проблемами є: порушення працездатності технічного та програмного забезпечення, поширення інформаційної зброї, постійне ускладнення інформаційних і комунікаційних систем, можливість концентрації інформаційних ресурсів у руках невеликої групи власників, шкодливе використання інформаційних даних, маніпулювання свідомістю, технологічний вплив на психічну діяльність.

Проте разом із цим технології штучного інтелекту розглядаються як одне з найефективніших засобів у сфері кібербезпеки як зараз, так і в майбутньому.

Виявлення шахрайства, виявлення шкідливих програм, виявлення вторгнень, оцінка ризиків у мережі та аналіз поведінки користувачів/машин – це п'ять найбільш актуальних способів використання штучного інтелекту (ШІ) для покращення кібербезпеки. ШІ дійсно змінює традиційні аспекти кібербезпеки. Він підвищує здатність компаній передбачати і запобігати кіберзлочинам, захищає пристрої з нульовим рівнем довіри і навіть може контролювати терміни дії паролів! Таким чином, штучний інтелект дійсно необхідний для забезпечення безпеки будь-яких об'єктів господарської або фінансової діяльності.

Пошук зв'язків між загрозами і аналіз шкідливих файлів, підозрілих IP-адрес або незвичної діяльності співробітників займає лічені секунди або хвилини. ШІ вже зараз допомагає людям забезпечувати кібербезпеку. А в майбутньому його можливості будуть тільки розширюватися, роблячи участь людини в процесі захисту суто номінальною.

У банках завдяки ШІ антифрод-системи стануть більш надійними та швидкими, що дозволить зекономити довіру і гроші як клієнтів фінансових установ, так і самих банкірів. На думку компанії Dell, яка займається розробкою таких продуктів, ШІ здатний захищати, контролювати та відстежувати дані в гібридних середовищах, а також запобігати 99% атак шкідливого ПЗ.

Крім того, ШІ може бути реалізований у вигляді хмари. Це дозволить йому автоматично масштабуватися при різкому збільшенні навантаження (наприклад, якщо хакери намагаються «атакувати» сервер або маскувати свою активність під лавину типових дій в іншому напрямку). Хмара дозволить розширити безпечний периметр компанії, якщо вся носима електроніка (гаджети) також буде підключена до контролюваного ШІ середовища.

Література

19. Неретін О., Харченко В. Забезпечення кібербезпеки систем штучного інтелекту: аналіз вразливостей, атак і контрзаходів. Information Systems And Networks. 2022. № 12. С. 7-20.

1. Савченко В.А., Шаповаленко О.Д. Основні напрями застосування технологій штучного інтелекту

у кібербезпеці. Сучасний захист інформації. 2020. № 4 (44). С. 6-11.

2. Стьопочкіна І.В., Новіков О.М. Методи штучного інтелекту в кібербезпеці: навч. посіб. для здобувачів спец. 125 “Кібербезпека”. Київ: КПІ ім. Ігоря Сікорського, 2022. 82 с.

3. Шаров С.В. Сучасний стан розвитку штучного інтелекту та напрямки його використання: зб. наук. пр. Інноваційні обрії України. 2023. № 6. С.136-144. – (Громадська організація Українські студії в європейському контексті).

4. Цяпа С.М. Огляд зарубіжних законодавчих ініціатив стратегічного використання технологій штучного інтелекту в сучасних умовах. Інформація і право. № 2(37)/2021. С. 51-59.

5. Tuomo Sipola, Tero Kokknen, Mika Karjalainen Artificial Intelligence and Cybersecurity: Theory and Applications. JAMK University of Applied Sciences. Publisher: Springer; 1st ed. 2023 edition (December 8, 2022). 311 p. DOI 10.1007/978-3-031-15030-2

6. Narcisa Roxana Mosteanu. Artifical Inelligence and cyber security – face to face with cyber attack – a maltese case of risk management approach. Ecoforum journal. 2020. Vol 9. № 2. URL: http://www.ecoforumjournal.ro/index.php/eco/article/view/1059

7. Rammanohar Das, Raghav Sandhane. Artificial Intelligence in Cyber Security. ICACSE 2020. IOP Publishing. Journal of Physics: Conference Series 1964 (2021). P.1-10 doi:10.1088/1742-6596/1964/4/042072. URL: https://iopscience.iop.org/article/10.1088/1742-6596/1964/4/042072/pdf

8. Гладка Ю.А., Назаренко Є.О. Аналіз застосування технологій штучного інтелекту в кібербезпеці: наукові праці третьої Міжнар. наук.-практ. конф. Сучасні тенденції розвитку інформаційних систем і телекомунікаційних технологій, м. Київ, 25 – 26 січня 2021 р. Київ: НУХТ, 2021. С. 64-66.

9. Про схвалення Концепції розвитку штучного інтелекту в Україні: Розпорядження Кабінету Міністрів України від 02.12.20 р. № 1556 URL: https://zakon.rada.gov.ua/laws/show/ 1556-2020-р#Text

10. Про затвердження Плану заходів з реалізації Концепції розвитку штучного інтелекту в Україні на 2021 – 2024 роки: Розпорядження Кабінету Міністрів України від 12.05.21 р. № 438 URL: https://zakon.rada.gov.ua/laws/show/438-2021-р#Text

11. Федоров: в Україні став доступний чат-бот зі штучним інтелектом ChatGPT. – (Українські національні новини від 18.02.23 р.). URL: https://www.unn.com.ua/uk/news/20160 33-fedorov-v-ukrayini-stav-dostupniy-chat-bot-zi-shtuchnim-intelektom-chatgpt

12. ChatGPT. The impact of Large Language Models on Law Enforcement. Europol Public Information. URL: https://www.europol.europa.eu/cms/sites/default/files/documents/Tech%20Watch%20Flash%20Enforcement.pdf

13. Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence. URL: https://ec.europa.eu/commission/presscorner/detail/en/IP_21 _1682.

References

1. Neretin, O., & Kharchenko, V. (2022). Ensuring cybersecurity for artificial intelligence systems: Analysis of vulnerabilities, attacks, and countermeasures. Information Systems and Networks, 12, 7-20.

2. Savchenko, V. A., & Shapovalienko, O. D. (2020). Main directions of applying artificial intelligence technologies in cybersecurity. Modern Information Protection, 4(44), 6-11.

3. Stiopochkina, I. V., & Novikov, O. M. (2022). Methods of artificial intelligence in cybersecurity: A textbook for students of specialty 125 “Cybersecurity”. Kyiv: KPI named after Igor Sikorsky.

4. Sharov, S. V. (2023). Current state of artificial intelligence development and its application directions. Innovative Horizons of Ukraine, 6, 136-144.

5. Tsapa, S. M. (2021). Review of foreign legislative initiatives for the strategic use of artificial intelligence technologies in modern conditions. Information and Law, 2(37), 51-59.

6. Sipola, T., Kokknen, T., & Karjalainen, M. (2022). Artificial Intelligence and Cybersecurity: Theory and Applications. Springer. https://doi.org/10.1007/978-3-031-15030-2

7. Mosteanu, N. R. (2020). Artificial Intelligence and Cybersecurity – Face to Face with Cyber Attack – A Maltese Case of Risk Management Approach. Ecoforum Journal, 9(2). http://www.ecoforumjournal.ro/index.php/eco/article/view/1059

8. Das, R., & Sandhane, R. (2021). Artificial Intelligence in Cyber Security. Journal of Physics: Conference Series, 1964, 1-10. https://doi.org/10.1088/1742-6596/1964/4/042072

9. Hladka, Y. A., & Nazarenko, Y. O. (2021). Analysis of the use of artificial intelligence technologies in cybersecurity. In Proceedings of the Third International Scientific and Practical Conference on Modern Trends in the Development of Information Systems and Telecommunication Technologies (pp. 64-66). Kyiv: NUHT.

10. Cabinet of Ministers of Ukraine. (2020). On Approval of the Concept for the Development of Artificial Intelligence in Ukraine: Order of December 2, 2020, No. 1556. https://zakon.rada.gov.ua/laws/show/1556-2020-p#Text

11. Cabinet of Ministers of Ukraine. (2021). On Approval of the Action Plan for Implementing the Concept for the Development of Artificial Intelligence in Ukraine for 2021–2024: Order of May 12, 2021, No. 438. https://zakon.rada.gov.ua/laws/show/438-2021-p#Text

12. Ukrainian National News. (2023). Fedorov: A Chatbot with Artificial Intelligence ChatGPT is now available in Ukraine. https://www.unn.com.ua/uk/news/2016033-fedorov-v-ukrayini-stav-dostupniy-chat-bot-zi-shtuchnim-intelektom-chatgpt

13. Europol. (n.d.). ChatGPT: The Impact of Large Language Models on Law Enforcement. https://www.europol.europa.eu/cms/sites/default/files/documents/Tech%20Watch%20Flash%20Enforcement.pdf

14. European Commission. (2021). Europe Fit for the Digital Age: Commission Proposes New Rules and Actions for Excellence and Trust in Artificial Intelligence. https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1682