Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)

Кафедра Інформаційно-мережна інженерія
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Інформаційні ризики використання штучного
інтелекту в інфокомунікаційних мережах
(тема)

Виконав:
здобувач 2 курсу, групи ІМІм-23-1

Кабаченко В.О.
(прізвище, ініціали)

Спеціальність 172 Електронні комунікації та радіотехніка
(код і повна назва спеціальності)

Тип програми освітньо-професійна
Освітня програма
Інформаційно-мережна інженерія
(повна назва освітньої програми)

Керівник доцент Золотарьов В.А.
(посада, прізвище, ініціали)

Допускається до захисту

зав. кафедри ІМІ

_____ Безрук В.М.
(підпис) (прізвище, ініціали)

2025р.

Не містить відомостей, заборонених до відкритого публікування

Кабаченко В.О

Золотарьов В.А.

Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
Кафедра Інформаційно-мережна інженерія
Рівень вищої освіти другий (магістерський)
Спеціальність 172 Електронні комунікації та радіотехніка
Тип програми Освітньо-професійна
Освітня програма Інформаційно-мережна інженерія
(шифр і назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри ІМІ _____
(підпис)

«»_____2025 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові *Кабаченку Вячеславу Олеговичу*

1. Тема роботи *Інформаційні ризики використання штучного інтелекту в інфокомунікаційних мережах*
Затверджена наказом по університету від 24 жовтня 2024 року, 1148 Ст

2. Термін подання студентом роботи 15 січня 2025 року

3. Вихідні дані до роботи: Проаналізувати інформаційні ризики використання систем штучного інтелекту в інфокомунікаційних системах, дослідити основні типи атак з використанням СШІ, проаналізувати загрози та атаки за рівнем тяжкості, оцінити ризики безпеки

4. Зміст пояснювальної записки: Вступ, 1. Правові проблеми використання штучного інтелекту 2. Загроза використання ші. 3.Оцінювання та порівняльний аналіз загроз штучного інтелекту. 4. Штучний інтелект як цифровий інструментарій забезпечення безпеки функціонування. 5. Аналіз атак із використанням ші. Висновки.

5. Перелік графічного матеріалу: Правові проблеми використання штучного інтелекту та юридичні обмеження використання ШІ в Україні та світі, загроза використання ШІ, ознаки які можуть свідчити про використання ШІ, порівняльний аналіз загроз ШІ, основні типи технологій штучного інтелекту в системі інформаційної безпеки, огляд атак на системи штучного інтелекту.

КАЛЕНДАРИЙ ПЛАН

№	Назва етапів кваліфікаційної роботи	Термін виконання етапів проєкту	Примітка
1.	Написання першого розділу	15.11.2024	Виконано
2.	Написання другого розділу	01.12.2024	Виконано
3.	Написання третього розділу	15.12.2024	Виконано
4.	Написання четвертого розділу	31.12.2024	Виконано
5.	Написання п’ятого розділу	10.01.2025	Виконано
6.	Оформлення пояснювальної записки	14.01.2025	Виконано
7.	Подання на перевірку	15.01.2025	Виконано

Дата видачі завдання 24 жовтня 2024 р.

Студент (підпис) В.О. Кабаченко

Керівник роботи (підпис) доцент Золотарьов В.А.

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: с.81, рис.8, табл.7,дод.2, джерел 38.

Об’єктом дослідження є штучний інтелект.

Мета роботи – дослідження найпопулярніших загроз штучного інтелекту.

В процесі дослідження проаналізовано основні види загроз, пов’язаних із застосуванням ШІ, включаючи атаки на інформаційну безпеку, маніпуляції з даними та вразливості алгоритмів навчання. Дослідження підкреслює важливість комплексного підходу до забезпечення безпеки в інфокомунікаційних мережах, включаючи поєднання традиційних засобів захисту з новітніми технологіями.

Висновки роботи можуть слугувати основою для подальших досліджень у сфері безпеки ШІ та інфокомунікацій.

ШТУЧНИЙ ІНТЕЛЕКТ; ЗАГРОЗА ШІ; ФІШИНГ; КІБЕРАТАКА; АНАЛІЗ ЗАГРОЗ.

ABSTRACT

Explanatory note of the qualification work: p.81, pic.8, tabl.7 ,app.2, sources 38.

The object of study is artificial intelligence.

The purpose of the work is to study the most popular threats to artificial intelligence.

The study analyzes the main types of threats associated with the use of AI, including attacks on information security security, data manipulation, and vulnerabilities of learning algorithms. The study emphasizes the importance of an integrated approach to ensuring security in infocommunication networks, including a combination of traditional security measures and with the latest technologies.

The conclusions of the paper can serve as a basis for further research in the field of AI security and infocommunications.

ARTIFICIAL INTELLIGENCE; CYBER THREAT; PHISHING; CYBER ATTACK; THREAT ANALYSIS.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ 9
ВСТУП 11
1 ПРАВОВІ ПРОБЕЛМИ ВИКОРИСТАННІ ШТУЧНОГО ІНТЕЛЕКТУ 14
1.1 Постановка задачі 14
1.2 Юридичні обмеження використанні ШІ в Україні та світі 15
1.3 Обмеження в Україні 16
1.4 Обмеження в Європейському Союзі та інших країнах 18
2 ЗАГРОЗА ВИКОРИСТАННЯ ШІ 24
2.1 Аналіз рішень 24
2.2 Загрози та ризики 28
2.3 Виявлення використання ChatGPT інструменти та потенціальні ризики 31
2.4 Основні ознаки які можуть свідчити про використання ШІ 32
3 ОЦІНЮВАННЯ ТА ПОРІВНЯЛЬНИЙ АНАЛІЗ ЗАГРОЗ ШТУЧНОГО ІНТЕЛЕКТУ 36
3.1 Аналіз загроз ШІ 36
3.2 Оцінка та реальні приклади загроз 38
4 ШТУЧНИЙ ІНТЕЛЕКТ ЯК ЦИФРОВИЙ ІНСТРУМЕНТАРІЙ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ФУНКЦІОНУВАННЯ 42
4.1 Перспективи застосування технологій штучного інтелекту 42
4.2 Використання штучного інтелекту в інформаційній безпеці держави 44
4.2.1 Endpoint Detection and Response 45
4.2.2 Network Detection and Response 47
4.2.3 User and Entity Behavior Analytics 48
4.2.4 Threat Intelligence Platform 49
4.2.5 Security Information and Event Management 49
4.2.6 Security Orchestration and Automated Response 51
4.2.7 Системи захисту застосунків 52
4.2.8 Антифрод 53
4.3 Застосування штучного інтелекту у сфері національної безпеки й обороноздатності держави 55
5 АНАЛІЗ АТАК ІЗ ВИКОРИСТАННЯМ ШІ 58

5.1 Постановка проблеми 58
5.2 Основні відомості про вразливості та класифікацію атак на системи штучного інтелекту 59
5.3 Огляд атак на системи штучного інтелекту 61
ВИСНОВКИ 66
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 68
ДОДАТОК А 74
ДОДАТОК Б 79

ПЕРЕЛІК УМОВНИХ ПОСИЛАНЬ

AGI (Artificial General Intelligence) – загальний штучний інтелект

AI (Artificial Intelligence) – штучний інтелект

CNIL (National Commission on Informatics and Liberty) – національна комісія з питань інформатики та свободи

EDR (Endpoint Detection and Response) – виявлення кінцевих точок та реагування

EXIF (Exchangeable Image File Format) – обмінний формат зображень

FCRA (Fair Credit Reporting Act) – акт про справедливе кредитування

GDPR (General Data Protection Regulation) – загальний регламент про захист даних

JS (JavaScript) – мова програмування

NDR (Network Detection and Response) – виявлення та реагування на мережі

PIPEDA (Personal Information Protection and Electronic Documents Act) – закон про приватність і захист електронних документів

SIEM (Security Information and Event Management) – інформація про безпеку та управління подіями

SOAR (Security Orchestration and Automated Response) – оркестрування безпеки та автоматизоване реагування

TIP (Threat Intelligence Platform) – платформа розвідки загроз

UEBA (User and Entity Behavior Analytics) – аналітика поведінки користувачів та організацій

ЄС – європейський союз

МН – машинне навчання

США – Сполучені Штати Америки

СШІ – система штучного інтелекту

МН – машинне навчання

ШІ – штучний інтелект

ВСПУП

Штучний інтелект, або скорочено ШІ — це галузь комп’ютерних наук, яка фокусується на розробці машин і систем, здатних виконувати завдання, що зазвичай вимагають людського інтелекту, такі як навчання, розв’язання проблем і прийняття рішень. В основі ШІ лежить ідея створення машин, які можуть мислити й міркувати, як люди, і можуть вчитися на власному досвіді, щоб з часом покращувати свою продуктивність. Сфера штучного інтелекту постійно розвивається і має потенціал революціонізувати багато аспектів нашого життя — від охорони здоров’я і фінансів до транспорту і розваг [23].

Державна безпека, національна безпека і оборона, а також багатосторонній розвиток суспільства залежать від розвитку високих технологій. Специфічною галуззю, що грає важливу роль в цьому процесі, стає штучний інтелект.

Принципи та алгоритми функціонування Систем Штучного Інтелекту переважної більшості суспільства практично невідомі. За суттю СШІ сприймаються як деякі «чарівні чорні скриньки», які здатні розуміти природну мову людини, музичні опуси або графічні зображення та адекватно реагувати на запитання користувачів шляхом надання статистично вірної відповіді. При цьому, звичайно, користувачі системи, що отримують результат відповідно до завдання, поставленого такій системі, не розуміють джерел формування відповіді і методів розв’язання завдання [6].

У сучасному світі технології штучного інтелекту та машинного навчання швидко розвиваються, знаходячи застосування в різних галузях, від медицини до фінансів, від виробництва до маркетингу. Водночас із цим розвитком виникає необхідність у вдосконаленні процесів розробки, розгортання та підтримки систем на основі ШІ та МН.

Потенційні можливості застосування штучного інтелекту величезні, і ми вже бачимо багато реальних прикладів того, як ця технологія використовується для покращення нашого життя.

- Персональні асистенти, такі як Siri та Alexa, використовують обробку природної мови, щоб розуміти наші запити й надавати інформацію або допомогу.

- Самокеровані автомобілі використовують комп'ютерний зір, щоб "бачити" дорогу попереду і приймати рішення про те, як нею рухатися.

- Системи виявлення шахрайства використовують машинне навчання для виявлення незвичайних шаблонів у фінансових даних, допомагаючи запобігти шахрайським діям.

- Медичні дослідження також отримують користь від ШІ: розробляються системи для аналізу даних про пацієнтів і прогнозування наслідків захворювань [23].

- У фінансовій сфері ШІ використовується для розробки моделей прогнозування цін на акції та інвестиційних можливостей.

- У маркетингу ШІ використовують для аналізу поведінки споживачів і персоналізації рекламних кампаній.

- У виробництві штучний інтелект використовується для оптимізації виробничих процесів і зменшення відходів.

- В освіті ШІ використовують для розробки персоналізованих навчальних програм і допомоги в оцінюванні.

Процеси автоматизації та самонавчання можуть призводити до неочікуваних наслідків, таких як вразливості у безпеці, масштабовані кібератаки або зловживання інформацією. У цьому контексті важливим є вивчення інформаційних ризиків, що виникають при впровадженні ШІ, а також розробка ефективних стратегій управління цими ризиками.

Особливо важливим є його застосування в інфокомунікаційних мережах, де ШІ забезпечує ефективну маршрутизацію даних, підвищує рівень безпеки та оптимізує комунікаційні процеси. Проте, разом з безперечними перевагами, використання ШІ супроводжується численними інформаційними ризиками, які можуть суттєво вплинути на цілісність, конфіденційність та доступність даних.

1 ПРАВОВІ ПРОБЛЕМИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ

1.1 Поставка задачі

З науково-технічного погляду штучний інтелект є сферою комп'ютерних наук, спрямованою на створення та розвиток систем здатних аналізувати інформацію, розпізнавати образи, розуміти мову та приймати рішення відповідно до певних критеріїв. З юридичного – визначення штучного інтелекту стає складнішим завданням. Враховуючи швидкий розвиток сфери штучного інтелекту, і правове регулювання повинно постійно еволюціонувати, щоби враховувати нові виклики та можливості. Штучний інтелект стрімко розвивається, і його використання охоплює дедалі більше сфер життя. Це відкриває безліч можливостей, але й породжує нові правові проблеми, які потребують вирішення.

Перелік правових проблем:

Відсутність чіткого правового визначення ШІ. В українському законодавстві наразі немає чіткого визначення ШІ, що ускладнює його правове регулювання.

Інтелектуальна власність. Не визначено, хто володіє авторськими правами на твори, створені ШІ-системами.

Відповідальність за використання ШІ. Виникає питання, хто несе відповідальність за шкоду, заподіяну діями ШІ-систем: розробники, власники, користувачі чи ніхто?

Безпека. Зловживання ШІ може становити загрозу для національної безпеки та кібербезпеки.

Вирішити ці проблеми, можемо наступними кроками, такими як, розробка законодавства, тобто законопроекту про ШІ, який має визначити його правовий статус та правила використання і міжнародне співробітництво, тобто Україна має співпрацювати з іншими країнами для розробки спільних підходів до регулювання ШІ.

1.2 Юридичні обмеження використання ШІ в Україні та світі

В українському законодавстві запровадження ШІ досі перебуває на етапі концепції та планування. Відомо, що Кабінет Міністрів України своїм схвалив Концепцію розвитку штучного інтелекту в Україні. ШІ планують включити до значної кількості сфер суспільного життя: від освіти і до правосуддя [9]. Пізніше, у 2021 році, Кабмін затвердив План заходів із реалізації цієї концепції на строк до 2024 року. Також, у жовтні 2023 року Міністерство цифрової трансформації оприлюднило дорожню карту з регулювання штучного інтелекту в Україні [10], що складається із двох етапів терміном в три роки, кінцевою метою якої є імплементація акту Європейського Союзу про штучний інтелект (AI Act), що в свою чергу свідчить про спроби гармонізації українського законодавства з правом ЄС [2].

Мета у ЄС – сприяти розвитку надійного етичного ШІ, який відповідає нормам законодавства. Із цією метою ЄС працює над адаптацією законодавства, розробляє етичні принципи та правові норми, що стосуються використання ШІ. Зокрема, починаючи з 2021 року в рамках ЄС розробляється перший у світі закон про ШІ. У його проекті вперше надано розгорнуте правове визначення ШІ в ЄС, яке може стати прикладом для інших країн, зокрема й України. Документ визначає, що система ШІ – це машина зі здатністю до навчання, яка розроблена для функціонування з різним рівнем автономності та може для явних чи

прихованих цілей створювати результати, такі як прогнози, рекомендації або рішення, що впливають на фізичне чи віртуальне середовище [1].

Існує декілька прикладів застосування норм GDPR європейськими регуляторами до розробників систем ШІ. Так, компанія Clearview AI – розробник програмного забезпечення для розпізнавання облич з використанням штучного інтелекту – отримала декілька штрафів через незаконне використання персональних даних осіб.

Управління із захисту персональних даних у Франції (CNIL), зокрема, що Clearview AI використовує особисту інформацію без законної згоди користувачів і водночас немає законного інтересу для такого збору, що є серйозним порушенням GDPR.

Італія обмежила доступ до всім відомого Chat GPT через занепокоєння щодо оброблення персональних даних громадян. Так, італійський регулятор зазначає про відсутність правової основи, яка б виправдовувала масовий збір і зберігання персональних даних із метою «навчання» алгоритмів, що лежать в основі роботи платформи [11].

1.3 Обмеження в Україні

Закон України “Про захист персональних даних”. Вимагає захисту персональних даних, що є особливо важливим у контексті ШІ, який може обробляти великі обсяги таких даних.

Основні положення:

Мета закону. Головною метою цього закону є захист прав і свобод людини та громадянина у сфері обробки персональних даних.

Визначення персональних даних. За українським законодавством, персональні дані – це будь-яка інформація, що стосується фізичної особи, яка може бути ідентифікована, зокрема ім’я, прізвище, адреса, електронна пошта, дані про здоров’я тощо.

Обробка персональних даних. Закон передбачає, що обробка персональних даних може здійснюватися лише за згодою суб’єкта даних або в інших випадках, визначених законодавством.

Обмеження на використання. Штучний інтелект може обробляти великі обсяги персональних даних, тому компанії та організації, що використовують такі технології, повинні дотримуватись вимог щодо захисту даних, зокрема забезпечити безпеку та конфіденційність інформації.

Важливість у контексті ШІ. При використанні ШІ для обробки даних, компанії повинні впроваджувати механізми для захисту даних, такі як псевдонімізація, анонімізація чи шифрування.

Необхідно дотримуватись принципу мінімізації даних, що означає, що слід збирати лише ті дані, які є істотними для досягнення конкретних цілей [12].

Закон України “Про інформацію”. Регулює використання інформації, включаючи дані, які можуть використовуватися для тренування ШІ.

Основні положення:

Мета закону. Регулювання відносин, пов’язаних із доступом до інформації, її поширенням та використанням.

Визначення інформації. Закон описує, що таке інформація і які види інформації можуть існувати, включаючи дані, які використовуються для наукових цілей, статистичного аналізу, а також для створення алгоритмів ШІ.

Права на інформацію. Громадяни мають право на доступ до інформації, яка знаходиться у володінні держави, а також до наукових і культурних матеріалів, зокрема в контексті даних, що можуть використовуватися для навчання ШІ.

Важливість у контексті ШІ. Використання даних для навчання ШІ підпадає під регуляцію цього закону, що передбачає обов'язок організацій дотримуватись уніфікованих норм про використання інформації.

В умовах збору та використання інформації, організації мають забезпечити, щоб ці дані не містили персональних даних без належної обробки і захисту.

Необхідність дотримання етики використання інформації: організації повинні враховувати етичні норми при використанні даних для розробки ШІ, щоб уникати дискримінації та забезпечити прозорість [13].

1.4 Обмеження в Європейському Союзі та інших країнах

Європейський закон про ШІ (EU AI Act). Пропонує регулювання використання ШІ на основі ризиків, пов'язаних з його застосуванням. Введено категорії ризиків (від низького до високого) і відповідні вимоги до кожної.

Основні положення:

Мета регулювання. Основною метою EU AI Act є забезпечення безпеки та основних прав людей, а також сприяння розвитку інновацій у сфері ШІ шляхом створення прозорого, надійного та етичного регулювання.

Ризикова категоризація. Закон класифікує застосування ШІ на основі рівнів ризиків, які воно може нести.

Низький ризик. Застосування, яке не є потенційно небезпечним для основних прав і безпеки. Наприклад, використання чат-ботів для обслуговування клієнтів.

Середній ризик. Застосування, яке потребує дотримання певних вимог безпеки та прозорості. Це може включати системи, які надають рекомендації в медицині або фінансах.

Високий ризик. Системи, які можуть суттєво вплинути на права і свободи людини, наприклад, ШІ для критичних інфраструктур (енергетика, транспорт), а також для продуктивності у правоохоронних органах. Для таких систем вводяться найсуворіші вимоги та контроль.

Вимоги для високого ризику. Необхідність проведення оцінки відповідності перед введенням системи в експлуатацію. Зобов'язання надавати прозорість, включаючи інформацію про алгоритми, дані та методології, які були використані.

Вимога щодо моніторингу та звітності, включаючи звіти про продуктивність і безпеку системи.

Загальний регламент захисту даних (GDPR). Вимагає захисту персональних даних і накладає обмеження на автоматизовану обробку даних і профілювання.

Основні положення:

Мета регламенту. GDPR покликаний захистити права одиничних осіб на контроль за своїми персональними даними в рамках ЄС, підвищити прозорість обробки даних і забезпечити правову відповідальність організацій.

Захист персональних даних. GDPR встановлює жорсткі правила щодо збору, обробки та зберігання персональних даних.

Суб'єкти даних (особи, щодо яких обробляються дані) отримують право знати, яким чином їхні дані обробляються, а також право на доступ, виправлення, видалення і обмеження обробки своїх даних.

Обмеження на автоматизовану обробку та профілювання. GDPR забороняє автоматизоване прийняття рішень, включаючи профілювання, яке має правові наслідки для осіб або суттєво на них впливає, якщо це не обґрунтовано згоду суб’єкта даних або не є законодавчою вимогою. У випадках, коли автоматизоване прийняття рішення можливе, особа має право вимагати людського втручання та оскаржувати рішення [14].

Обмеження в США.

Законодавство про конфіденційність даних має на меті захист персональних даних осіб у різних сферах, включаючи бізнес, охорону здоров'я та технології. Основні принципи цього закону можна узагальнити наступним чином:

Збирання даних. Організації повинні обґрунтовано інформувати користувачів про мету збору їхніх персональних даних. Необхідно отримати згоду користувача перед їхнім збором і обробкою.

Обробка даних. Дані повинні бути оброблені законно, коректно та прозоро. Використання даних обмежується тими цілями, для яких вони були зібрані, і не можуть використовуватися для інших цілей без додаткової згоди.

Захист даних. Організації зобов’язані вживати заходів безпеки для захисту персональних даних від несанкціонованого доступу, втрати або знищення. Передбачаються штрафи за порушення, які можуть включати компенсацію постраждалим особам.

Права суб'єктів даних. Люди мають право на доступ до своїх персональних даних, їх коригування та видалення. Закони можуть включати положення про право на заперечення проти обробки даних у певних випадках.

Акт про справедливе кредитування. Цей закон регулює використання інформації про кредитоспроможність, надаючи певні права споживачам в контексті отримання кредиту.

Основні аспекти включають:

Регулювання кредитних бюро. Закон вимагає, щоб кредитні бюро діяли відповідно до певних стандартів точності та прозорості стосовно інформації, яку вони збирають та поширюють. Споживачі мають право на безкоштовний доступ до своїх кредитних звітів щорічно.

Використання ШІ в оцінці кредитоспроможності. Акт вимагає, щоб усі критерії, що використовуються для оцінки кредитоспроможності (в тому числі алгоритми ШІ), були обґрунтованими і не дискримінували конкретні групи населення. Під час використання автоматизованих систем, таких як ШІ, для оцінки кредитоспроможності, необхідно забезпечити, щоб дані, які аналізуються, були актуальні та репрезентативні.

Права споживачів. Споживачі мають право коригувати неточності у своїх кредитних звітах. У разі відмови у кредиті на основі інформації з кредитного звіту, споживач має право отримати копію звіту і дані про агентство, яке його надало [15].

Обмеження в Канаді

Закон про приватність і захист електронних документів. Канада має один із найсуворіших законодавчих актів, що захищає персональні дані, включаючи ті, які обробляються з допомогою ШІ.

Ключові аспекти:

Прозорість. Організації зобов'язані повідомляти, як і з якою метою вони збирають, використовують та розкривають особисті дані.

Право на доступ. Користувачі мають право отримати інформацію про свої дані, що зберігаються, і вимагати їх видалення за певних обставин.

Етичні стандарти. Закон також заохочує дотримання етичних норм під час обробки даних, що важливо для впровадження алгоритмів ШІ, які можуть впливати на особисте життя людей [17].

Таблиця 1.1 – Основний порівняльний аналіз обмежень

Регламент / Країна	**Основні аспекти**	**Обробка даних**	**Захист даних**	**Права суб'єктів даних**	**Обмеження**
США	Регулювання кредитних бюро.	Доступ до особистої інформації.	Необхідність захисту.	Право доступу до інформації.	Обробка на основі згоди.
Канада	Захист електронних документів.	Закон повинен забезпечувати прозорість.	Організації повинні захищати дані.	Право на доступ до своїх даних.	Обробка даних лише за згодою.
ЄС (GDPR)	Захист прав осіб на контроль.	Збір та обробка даних з обмеженнями.	Жорсткі правила щодо обробки.	Право на видалення та обмеження обробки даних.	Заборона автоматизованого прийняття рішень.
ЄС (EU AI Act)	Регулювання використання ІШ.	Класифікація ризиків.	Безпека прав людей.	Вимоги до прозорості алгоритмів.	Вимоги до моніторингу систем.
Україна ("Про інформацію")	Визначення інформації.	Регулювання доступу та використання даних.	Необхідність дотримання принципу мінімізації.	Право доступу до інформації.	Вимога про урахування етики.

Продовження таблиці 1.1

Україна ("Про захист персональних даних")	Захист прав людей.	Обробка даних лише за згодою.	Вимога захисту даних.	Право на доступ та контроль.	Заборона на обробку без згоди.

2 ЗАГРОЗИ ВИКОРИСТАННЯ ШІ

2.1 Аналіз рішень

Штучний інтелект (ШІ) – це потужна технологія, яка може принести багато користі суспільству. Однак важливо усвідомлювати й потенційні загрози, пов’язані з його використанням. Звідси випливають і безпекові ризики: мовні моделі дуже зручно використовувати для масової дезінформації, причому персоналізованої дезінформації «з людським обличчям», заточеної під конкретні групи людей, наприклад, шляхом коментування у соцмережах та у медіа. Це також може бути персоналізоване шахрайство, коли людина отримує повідомлення з урахуванням методів соціальної інженерії, які спрямовані саме на неї, які використовують індивідуальні слабкості та упередження — таке шахрайство буде краще працювати. Звісно, є також ризики, які стосуються порушення прав інтелектуальної власності та плагіату [2].

Окремі застосування ШІ, як-от дистанційна біометрична ідентифікація - технології ідентифікації за обличчям, ходою, поставою людини тощо, будуть високо ризиковими, бо можуть порушувати права людини або загрожувати її безпеці та здоров'ю. Також можна використовувати мовні моделі для допомоги у розробленні тих же комп’ютерних вірусів або зброї, коли штучним інтелектом буде зміщуватися певна експертиза, відсутня у людей, що діють зі злими намірами [4].

Використання символу «Ш» може бути загрозою або створювати небезпеку в декількох контекстах:

Фальшиві документи. Символ «Ш» може використовуватися для створення підроблених документів або підписів, які виглядають автентично.

Фішинг. У кібератаках символ «Ш» може використовуватися в URL адресах або електронних листах для введення користувачів в оману та змушування їх перейти на шкідливі вебсайти.

Кібератаки. Штучний інтелект (ШІ) відіграє важливу роль він спочатку використовував прості правила для відстеження мережевого трафіка та дій користувачів, а потім відбувається атака, яка заповнює вашу мережу фальшивими запитами, щоб порушити ваші бізнес-операції, яка називається DoS [24].

Шифрування даних. Зловмисники можуть використовувати символ «Ш» або подібні символи в алгоритмах шифрування для створення складних та важко зашифрованих повідомлень.

Маніпуляція текстом. Введення символу «Ш» в тексти або коди може змінювати зміст або функціональність програм, що може призвести до непередбачених наслідків.

Соціальні мережі та комунікації. Використання символу «Ш» в контекстах, де він може бути сприйнятий як загрозливий або агресивний, може створювати конфлікти та непорозуміння.

Виявлення дипфейків і підроблених фотографій, створених за допомогою штучного інтелекту, є важливим завданням у сучасному світі цифрових технологій. Нижче наведено основні методи та інструменти для виявлення дипфейків, їх характеристики, принципи дії, а також недоліки та переваги.

Алгоритми дипфейків на основі машинного навчання. Принцип дії. Ці алгоритми навчаються розпізнавати дипфейки на основі аналізу великої кількості реальних і підроблених зображень. Вони використовують різноманітні функції, такі як текстури, кольорові палітри та аномалії, що характерні для штучно змінених зображень.

Переваги. Можливість автоматизованого аналізу великих обсягів даних, можуть досягати високої точності в залежності від якості навчання.

Недоліки. Вимагають великих обсягів навчальних даних, можуть бути неефективними проти нових технік дипфейків.

Форензичний аналіз зображень. Принцип дії. Цей метод використовує численні техніки для оцінки зображень на основі їх фізичних характеристик. Форензичний аналіз може виявляти артефакти компресії, аномалії в освітленні та інші особливості, які вказують на те, що зображення було змінено.

Переваги. Не потребує попередньої інформації про зображення та може виявляти не тільки дипфейки, а й інші типи підробок.

Недоліки. Потребує спеціалізованих знань для інтерпретації результатів та іноді може давати помилкові тривоги через природні артефакти зображень.

Методи блокчейн-технологій. Принцип дії. Блокчейн може бути використано для перевірки автентичності зображень, записуючи їх метадані під час створення. У будь-який момент можна перевірити, чи є зображення оригінальним.

Переваги. Високий рівень безпеки та довіри до даних, можливість стежити за історією зображення.

Недоліки. Вимагає інфраструктури та навчання користувачів та може бути неефективним для вже поширених зображень без зафіксованої автентичності.

Методи аналізу метаданих. Принцип дії. Це техніка, яка аналізує метадані зображень, такі як EXIF, щоб виявити зміни або аномалії, які можуть вказувати на маніпуляції.

Переваги. Простота використання та доступність і можливість бути швидким способом виявлення підробок.

Недоліки. Метадані можуть бути легко змінені або видалені та не завжди відображає істину про вміст зображення.

Системи виявлення на основі нейронних мереж. Принцип дії. Використання глибинних нейронних мереж для детекції дипфейків шляхом аналізу пікселів зображення на різних рівнях.

Переваги. Можуть виявляти складні моделі підробок і постійно удосконалюються завдяки новим дослідженням.

Недоліки. Необхідні потужні обчислювальні ресурси які можуть бути уразливими до нових алгоритмів дипфейків та перехитрують навчені моделі.

Одним із найпоширеніших ризиків, пов’язаних із штучним інтелектом (ШІ), є можливість втратити контроль над ним та його надзвичайними здібностями. Експерти по штучному інтелекту (ШІ) вважають, що його небезпечність можна порівняти з ядерною війною чи глобальною пандемією. Вони переконані, що ми швидко рухаються в напрямку загибелі людства через ШІ, і що введення регулювання є обов’язковим вже зараз. Інші ж вважають, що ці занепокоєння є занадто гіперболічними. Вони стверджують, що ШІ - це інструмент, який може бути використаний як на благо, так і на зло, і що важливо, щоб люди контролювали його розвиток.

Потенційні ризики ШІ, що представили дослідники.

Втрата контролю над ШІ. Одним із найпоширеніших ризиків, пов’язаних із штучним інтелектом (ШІ), є можливість втратити контроль над ним та його надзвичайними здібностями. Загальний штучний інтелект (AGI) є однією з головних загроз, оскільки він може бути таким же розумним або навіть розумнішим за людей у широкому спектрі завдань.

ШІ та безробіття. Втрата робочих місць через автоматизацію є серйозною загрозою для багатьох. Деякі керівники розповідають про свої плани щодо використання ШІ відкрито, і це може призвести до масового безробіття. Спеціалісти відзначають, що ця загроза може призвести до втрати сенсу життя для людей, які втратять свої робочі місця.

Упередженість. Упередженість може виникнути внаслідок різних факторів. Наприклад, якщо навчальні дані містять лише образи людей білої раси, то системи ШІ, навчені на цих даних, можуть бути більш схильними до помилок у розпізнаванні облич людей інших рас [5] .

2.2 Загрози та ризики

Одним із поштовхів до стрімкої діджиталізації у свій час став локдаун та популяризація віддаленого формату роботи. Оскільки ми дуже швидко звикаємо до автоматизованих процесів, ми прагнемо їх удосконалення, що пояснює вагомі розробки у сфері штучного інтелекту за останні роки. Вони уже змінюють окремі галузі та поступово оновлюють чи замінюють певні професії. Варто зазначити, що штучний інтелект може виконувати майже всю рутинну роботу за умови проведення відповідної реорганізації бізнес-процесів. Генеративні технології автоматизовані виконання таких завдань, як створення контенту, дизайн та обслуговування клієнтів, тим самим змінюючи робочі процеси. Це частково може призвести до втрати певних позицій або появи нових, таких як, наприклад, оператор 3D-принтера [8].

Розвиток технологій штучного інтелекту сприятиме появі нових можливостей працевлаштування за такими напрямками, як розробка штучного інтелекту, машинне навчання, аналіз даних та інші. Штучний інтелект має безліч переваг для його використання у роботі і повсякденному житті. Він значно оптимізує наші щоденні процеси написання електронних листів, повідомлень або маркетингових презентацій. Спеціалісти можуть тепер дуже швидко за потреби готувати статті або публікації на конференції із застосуванням ChatGPT. Як правило, технічним спеціалістам непросто формувати свої думки на папері, а штучний інтелект ефективно допомагає їм у цьому. Терміни для підготовки тих чи інших матеріалів скоротилися майже вдвічі, оскільки ChatGPT, для прикладу,

презентує готові блоки текстів та тематичні картинки за запитом. Його використання з преміальним акаунтом надає ще більше можливостей для пришвидшення виконання завдань як для окремих осіб, так і для компаній.

Пошук подібної інформації у браузері раніше займав значно більше часу. Однак, важливо пам’ятати, що будь-яку інформацію, отриману за допомогою штучного інтелекту, потрібно перевіряти. Спікери вебінару радять працювати із технологіями ШІ у діалоговому режимі для того, щоб структурувати отриману інформацію і, зокрема, мати можливість задавати питання тому чи іншому чат-боту. Наразі чат-боти, зокрема, ChatGPT можна вбудовувати у різні застосунки. Кількість додатків з таким функціоналом лише буде зростати і надавати все більше гарних можливостей для користувачів. Також поступово у таких додатках з’являться нові або будуть покращені існуючі текстові і голосові інтерфейси.

Одним із гарних прикладів уже розширеного функціоналу чат-боту є чат-бот Приват24, який постійно оновлюється і покращується. Окрім цього, варто зазначити, що наразі компаніям не потрібно розробляти окремі чат-боти, а достатньо лише придбати готовий сценарій JavaScript та застосувати його на сайті. Це забезпечить наступний рівень використання технологій кінцевими користувачами та підвищить привабливість тих чи інших компаній [7].

Розвиток технологій штучного інтелекту також приносить велику користь для ІТ-спеціалістів у їх роботі та розвитку ІТ-сфери у цілому. Популяризація ChatGPT та його вміння писати код передбачає, що безпосередня якість коду буде покращувати, швидкість його написання стане вищою, а процес управління кодом стане більш автоматизованим. Застосування технологій ШІ в ІТ-процесах удосконалює процеси програмування та управління проектами. Штучний інтелект поступово буде генерувати код на рівні досвідчених ІТ-спеціалістів рівнів junior та middle, уникаючи помилок. Роль же самих junior та middle фахівців поступово зміниться. Спікери вебінару вважають, що з часом ефективно працюватимуть нові команди, у яких весь проект зможуть будувати один

спеціаліст рівня senior та згодом один лідер команди, застосовуючи кілька спеціалізованих моделей, створених під спеціальні домени. Однак, на думку спікерів, класичні проекти з можливістю росту спеціалістів до рівня senior теж залишаться. На них junior та middle фахівці будуть застосовувати ШІ для вирішення тих чи інших завдань. З часом зміниться також і функціонал розробників DevOps, оскільки процеси хостингу та побудови контейнерів ставатимуть більш автоматизованими. Рішення з розробки будуть приймати безпосередньо IT-спеціалісти, однак усі процеси будуть значно пришвидшені та оптимізовані. Крім цього, штучний інтелект також полегшить роботу проектного менеджера, оскільки він зможе генерувати команду автоматично, спираючись на вимоги та цілі проекту.

Особливість штучного інтелекту та, зокрема, ChatGPT полягає у тому, що він базується в основному на двох процесах: обробці природньої мови (NLP) з використанням нейронних мереж для навчання та застосуванні моделі розподіленої обробки даних. Текст, згенерований ШІ, є унікальним і не повторюється. На основі відповідного запиту ШІ пропонує декілька варіантів, які ми можемо протестувати і обрати той, який підходить найкраще. Так, наприклад, NASA наразі активно застосовує штучний інтелект під час будівництва нових конструкцій, задаючи технічні параметри та завдання для подальшого генерування рішень. Технології штучного інтелекту дозволяють створювати футуристичні конструкції, які людина навіть і не змогла би колись придумати. Окрім цього, ШІ є дуже помічним у фінансовій сфері, зокрема, під час розрахунку біржових курсів, планування фінансових прогнозів та варіантів їх впливу на діяльність інвесторів. Штучний інтелект дозволяє автоматизувати багато важливих процесів, наприклад, і у гірничодобувній сфері. Технології ШІ також активно застосовуються у процесі моделювання моторизованих протезів за допомогою використання нейронних мереж, що людство не могло би уявити ще декілька десятків років тому. Штучний інтелект уже спромігся перевести

значну кількість наших повсякденних завдань у цифрову сферу, тим самим підвищивши нашу професійну ефективність.

Зрозуміло, що для того, щоб мати успіх і розвиватися, зокрема, у IT- сфері, спеціалісти повинні слідкувати за останніми трендами, навчатися та одночасно застосовувати нові технології. Саме за допомогою штучного інтелекту розробники тепер можуть виконувати свою роботу швидше і краще, оскільки алгоритмічні процеси за них уже може виконувати штучний інтелект. Однак, варто з обережністю підходити до питання надання доступу до особистої та корпоративної інформації, оскільки існують загрози зловживання такими даними. Наразі питання забезпечення їх захисту та збереження є одним із ключових і надзвичайно важливим для того, щоб продовжувати розвиток технологій ШІ безпечним чином та без завдання шкоди користувачам [3].

2.3 Виявлення використання ChatGPT інструменти та потенціальні ризики

З моменту запуску в листопаді 2022 року жоден додаток не мав більшого впливу на світ, ніж ChatGPT. Він майже одноосібно спричинив революцію в галузі штучного інтелекту, яка прокладає собі шлях майже в кожній галузі на Землі. На вершині цього списку - світ створення та публікації контенту. Тисячоліттями люди кладуть перо на папір або пальці на клавіатуру і вручну виводять слова на друк. Але з появою ChatGPT все змінилося назавжди. За допомогою кількох простих підказок люди можуть використовувати ChatGPT для створення статей за лічені секунди. Але з більш ніж 100 мільйонами користувачів по всьому світу, люди почали задавати питання, як виявити ChatGPT при читанні контенту в Інтернеті. Тож чи можна виявити ChatGPT? Безумовно [18].

По суті, жоден детектор, як людський, так і на базі штучного інтелекту, не може з точністю 100% визначити ChatGPT. Навіть якщо ви скористаєтеся ШІ-

детектором для аналізу конкретного тексту, він лише надасть ймовірність того, що він був створений штучним інтелектом. Детектор не вкаже, чи це зробив саме ChatGPT, чи, наприклад, Claude AI або Google Bard. Текст, написаний штучним інтелектом, може містити певні впізнавані патерни. Ці патерни проявляються у синтаксисі та структурі речень. Серед інших вимірів є розривність тексту, яка відображає варіацію довжини речень, а також розгубленість, що показує складність тексту. Кожен, хто працює з публікованими матеріалами чи контентом, має звертатися до інструментів для виявлення ChatGPT. Ці засоби особливо корисні для вчителів, які виявляють академічну нечесність, а також для редакторів, які стикаються з плагіатом, створеним штучним інтелектом. Проте ШІ-детектори можуть бути корисними для будь-кого, хто читає статті або блоги в інтернеті. Одна з проблем, пов’язаних з LLM, такими як ChatGPT, полягає в ризику отримання плагіатських або неправдивих результатів. Це явище відоме як галюцинації штучного інтелекту, і якщо ви не приділяєте уваги дослідженню та перевірці фактів, ви можете несвідомо поширити безперечно неправдиву інформацію.

2.4 Основні ознаки які можуть свідчити про використання ШІ

Виявлення закономірності та невідповідності. ШІ часто генерує тексти, які слідують певним моделям або шаблонам. Це може проявлятися у повторюваних фразах чи структурах речень. Якщо ви помічаєте, що текст має однаковий стиль або повторюється в різних частинах, це може бути ознакою автоматичного генерування.

Перевірка на наявність ознак людської помилки. Люди схильні робити певні типи помилок (граматичні, пунктуаційні, або стилістичні). Тексти, згенеровані ШІ, можуть містити менш поширені помилки або взагалі їх не мати.

Наприклад, якщо текст дуже «ідеальний» і не містить жодної помилки, це може бути підозрілим.

Звертайте увагу на помилки у мові. ШІ може використовувати менш емоційно насичену та образну мову, забуваючи про нюанси, які надають глибини писемному висловлюванню. Якщо текст відчувається сухим або безособовим, це може свідчити про його штучне походження.

Звертайте увагу на контекст. Штучний інтелект може не завжди вловлювати контекст певної ситуації або теми. Якщо текст відхиляється від основної теми або виглядає неприродно в контексті, це може бути показником його генерування ШІ.

Надмірне використання перехідних слів. ШІ іноді може зловживати перехідними словами для покращення зв'язності тексту, замість того аби використовувати їх стратегічно. Якщо ви помічаєте, що слова «однак», «далі», «також» використовуються занадто часто, це може свідчити про ШІ.

Речення, які виглядають правильно, але не мають сенсу. Інколи ШІ може створювати речення, граматично коректні, але семантично пусти. Такий текст може бути важким для розуміння або не відповідати логічному змісту.

Брак оригінальності. ШІ часто апелює до стереотипних ідей і фраз, не створюючи справді нових концепцій. Якщо текст здається передбачуваним або використовує кліше, це може вказувати на його генерування ШІ.

Фактичні помилки. Штучний інтелект може помилково подати фактичну інформацію через неправильне розуміння контексту або терміна. Регулярна перевірка фактів може виявити такі помилки, до яких людина не звернула б уваги [20].

Питання виявлення ChatGPT та подібних програм стає дедалі актуальнішим. Незважаючи на всі переваги цього інструмента, він далеко не досконалий. Університети та школи підвищують рівень контролю за плагіатом і

протидії зловживанню технологіями штучного інтелекту. Щоб зменшити ризики, важливо діяти проактивно. Вам слід бути обізнаними в методах виявлення. Користувачі повинні використовувати програми з штучним інтелектом з обережністю. Нижче ми надамо кілька стратегій, які можуть допомогти обійти виявлення ChatGPT, щоб залишитися непоміченими. Ці стратегії не є простими способами уникнення відповідальності, а, скоріше, рекомендаціями для розумного та етичного використання штучного інтелекту. Важливо розуміти як можливості, так і обмеження цього інструмента. Мета полягає в тому, щоб скористатися функціоналом ChatGPT, зберігаючи академічну та професійну доброчесність.

Неправильне використання ChatGPT може викликати серйозні проблеми, особливо в контексті роботи з конфіденційними даними. Недостатня обережність робить користувачів більш уразливими до витоку інформації. Робота з конфіденційною інформацією вимагає обачності. Ігнорування заходів безпеки може призвести до порушення конфіденційності даних. Безпека є першочерговою. Забезпечення безпеки даних має величезне значення. Ви повинні вжити необхідних заходів, щоб уникнути труднощів під час використання ChatGPT. Користувачі повинні бути уважними, аби запобігти витоку своїх даних [19].

Якщо студента спіймають на використанні ChatGPT, це може призвести до серйозних академічних наслідків, які вплинуть на його майбутнє. Можливі покарання варіюються від низьких оцінок до суворих дисциплінарних заходів,

включаючи відсторонення від навчання. Крім формальних санкцій, під загрозою опиняється репутація студента в академічному середовищі та на ринку праці. Він може втратити довіру своїх колег. Такі негативні наслідки можуть суттєво підшкодити кар'єрі та звести нанівець зусилля і ресурси, витрачені на професійну освіту.

Додатково до згаданих академічних ризиків, зловживання ChatGPT також може спричинити серйозні юридичні наслідки, зокрема можливі порушення авторських прав та інтелектуальної власності. Це може призвести до судових позовів, особливо якщо контент використовується в комерційних чи академічних цілях. Судові позови можуть негативно вплинути на довіру та репутацію, а також спричинити великі витрати, включаючи гонорари адвокатів та значні зусилля для виправлення помилок, навіть якщо вони сталися ненавмисно [19].

Отже, якщо контент, створений штучним інтелектом, не отримує високих оцінок, як можна зробити так, щоб детектори AI та навіть Google не виявили використання ChatGPT для його створення? Створити його невизначуваним. Undetectable.AI — це провідний інструмент для скремблювання AI, який перетворює ваш контент, написаний людиною чи штучним інтелектом, одним натисканням кнопки, на більш "людяний" варіант. Цей інструмент забезпечує те, що ваш перероблений контент не буде помічений як створений штучним інтелектом найкращими детекторами, такими як ZeroGPT, Writer, Copyleaks і Sapling. Незалежно від того, чи використовуєте ви ChatGPT для написання, чи самі створюєте контент, Undetectable.AI стане незамінним інструментом для кожного творця контенту, який хоче масштабувати свою роботу за допомогою штучного інтелекту. За 9,99 долара на місяць за 10 000 слів або лише 5,00 доларів на місяць при річній підписці, Undetectable.AI — це мінімальні витрати за душевний спокій будь-якого письменника.

3 ОЦІНЮВАННЯ ТА ПОРІВНЯЛЬНИЙ АНАЛІЗ ЗАГРОЗ ШТУЧНОГО ІНЕЛЕКТУ

3.1 Аналіз загроз ШІ

В цьому розділі буде проведено аналіз загроз безпеки ШІ, які нести велику загрозу. При створенні порівняльної таблиці будуть використовуватися дані, які вдалося знайти в офіційних джерелах і підтвердити. Якщо функція підтверджується, але є певні нюанси або особливості, що впливають на рівень безпеки.

При створенні таблиці загроз будуть перевірені такі загрози:

1. Фішингові атаки – це спроба отримати вашу особисту інформацію оманливим шляхом, використання ШІ для створення реалістичних листів та сайтів, що імітують достовірні джерела [21].

2. Розкриття інформації без змін – це ненавмисне або шахрайське розкриття конфіденційної інформації без її зміни (наприклад, витоки).

3. Підвищення обізнаності – ШІ використовується для навчання користувачів розпізнавати фішингові атаки та інші загрози.

4. Підміна співбесідника – це зловмисники видають себе за інших осіб у спілкуванні, щоб отримати конфіденційні дані.

5. Автоматизація шкідливого ПЗ – це використання ШІ для створення та розповсюдження шкідливого програмного забезпечення.

6. Виявлення фішингу за допомогою ШІ – може допомогти у виявленні фішингових атак через аналіз аномальних поведінок у мережі [21].

7. Модифікація інформації – це вихідне повідомлення змінюється або повністю підміняється іншим і надсилається адресату, це може призвести до дезінформації або фінансових збитків [22].

Спочатку проведемо порівняння загроз, проставимо відповідні оцінки та заповнимо у вигляді таблиці 1.

Таблиця 3.1.1 – Порівняння загроз ШІ за шкалою «Небезпечно – Безпечно»

Назва загрози	**Оцінка**
Фішингові атаки	Небезпечно
Розкриття інформації без змін	Небезпечно
Підвищення обізнаності	Безпечно
Підміна співбесідника	Небезпечно
Автоматизація шкідливого ПЗ	Небезпечно
Виявлення фішингу	Безпечно
Модифікація інформації	Небезпечно

Для того щоб провести розрахунок порівняння між загрозами ШІ застосуємо перехід зі шкали оцінювання «Небезпечно – Безпечно» до числової шкали оцінювання, наведеної в таблиці 3.2.

Таблиця 3.1.2 – Конвертація шкали оцінювання «Небезпечно – Безпечно»

Небезпечно	5
Безпечно	2

3.2 Оцінка та реальні приклади загроз

Для того щоб розуміти реальну небезпеку цих загроз, потрібно проаналізувати кожну з них і зрозуміти, на скільки та чи інша загроза буде нести за собою негативні чи позитивні наслідки. Перше що є важлим – це остерігатися підозрілої інформації та завжди перевіряти її, навіть якщо надіслали її із правдивого джерела. Перша загроза – це фішингові атаки, зазвичай ця атака супроводжується «оригінальними» сторінками, або застосунками, які на перший погляд не здаються підозрілими, але як тільки натиснути на посилання, або ввести свої дані і натиснути «Продовжити», то зазвичай після цього ваші дані надійдуть до рук 3-ї особи, яка і розпочала цю атаку. Для прикладу це може бути «Приват24» (рис. 3.1) , зазвичай зловмисник копіює даний сайт і змінює лиш посилання, але так, щоб на перший погляд взагалі не можна було відрізнити (рис. 3.2). На вигляд сайти ідентичні, але різниця в самому посиланні, яку взагалі не можна відрізнити, літера «p» (англійська) змінена на літеру «р» (українську).

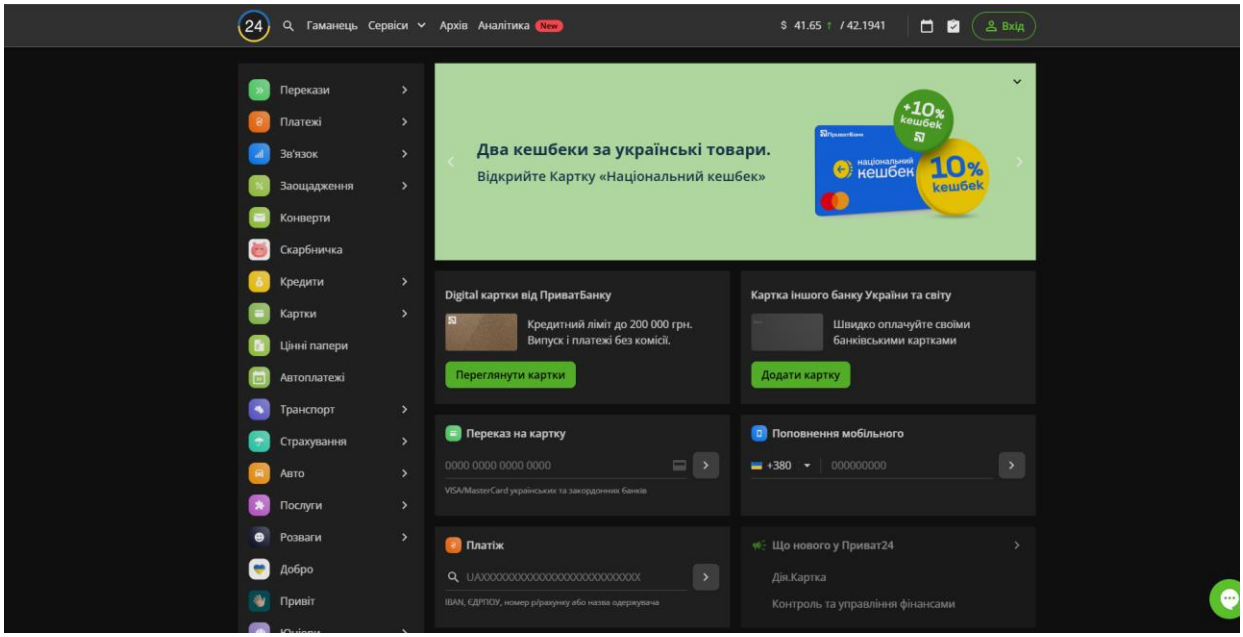

Рисунок 3.2.1 – Приват 24 https://next.privat24.ua/

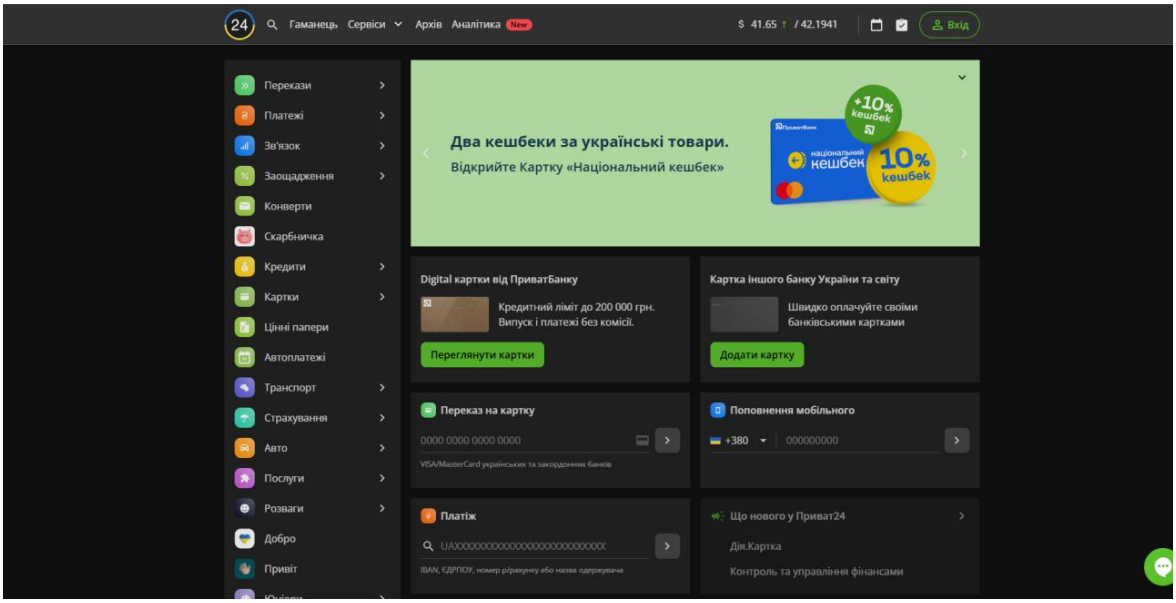

Рисунок 3.2.2 – Сайт зловмисника https://next.privat24.ua/

Друга загроза – це модифікація інформації та підміна співбесідника, це дві загрози, але їх можна поєднати в одну. Для прикладу модифікація інформації може відбуватися і без вашого відому, якщо хтось із довіреної вами особи, має доступ до вашої електронної пошти. Ця довірена особа може зайти на вашу

електронну пошту, завантажити необхідну інформацію та надіслати тій самій людині, але уже свою модифіковану із словами вибачення, що помилився при написанні того, чи іншого документа. Наступна загроза – це розкриття інформації без змін, для того щоб ця загроза минула вас стороною, то бажано використовувати шифрування даних та засоби контролю доступу. Щоб завжди було обмеження можливостей користувачів на доступ до чутливої інформації.

Зазвичай під час кібератак, використовують автоматизацію шкідливого ПЗ, що може призвести до збою всієї системи та може і взагалі до її припинення існування. Алгоритм цієї атаки починається із фішингу, який супроводжує за собою персоналізовані цілі та контент, що підвищує ймовірність успішного збору особистої інформації, потім йде перевірка а вразливість системи, що забезпечує виявити слабкі місця, а уже потім йде процес «зараження» системи шкідливим ПЗ. Отже проаналізувавши загрози, можна зробити детальну таблицю аналізу загроз та наслідків.

Таблиця 3.2.1 – Порівняльний аналіз

Загроза	**Ступінь небезпеки**	**Ймовірність реалізації**	**Наслідки**
Фішингові атаки	Висока	Висока	Витік даних, фінансові втрати
Розкриття інформації	Середня	Середня	Репутаційні втрати, компрометація
Підміна співбесідника	Висока	Середня	Шахрайство, викрадення критичних даних

Продовження таблиці 3.2.1

Автоматизація шкідливого ПЗ	Висока	Висока	Злом, пошкодження систем
Модифікація інформації	Висока	Середня	Втрата довіри, неправильні рішення

4 ШТУЧНИЙ ІНТЕЛЕКТ ЯК ЦИФРОВИЙ ІНСТРУМЕНТАРІЙ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ФУНКЦІОНУВАННЯ

4.1 Перспективи застосування технологій штучного інтелекту

Сьогодні штучний інтелект широко використовується в бізнесі як потужний інструмент для розвитку компаній. Він забезпечує безперервну взаємодію з клієнтами 24/7, удосконалює процеси електронної комерції, оптимізує роботу банків та інших ключових галузей економіки. Масова популярність ШІ розпочалася з запуску чат-бота ChatGPT від OpenAI у листопаді 2022 року. Ця мовна модель одразу стала популярною серед користувачів, які почали делегувати їй частину рутинних завдань, суттєво спрощуючи робочі процеси. ChatGPT, що використовує ШІ для створення тексту, коду чи відповідей на запити, став найбільш швидкозростаючим інтернет-додатком в історії. ШІ застосовується в різних сферах, таких як безпілотні автомобілі, чат-боти, розпізнавання облич, голосові асистенти (Alexa, Google Assistant, Siri), а також у рекламі на основі ШІ. За даними Frost & Sullivan, технології штучного інтелекту до 2030 року зможуть збільшити глобальний оборот компаній на 15,7 трильйона доларів, з яких 10,7 трильйона припадатиме на США та Китай. Подібні прогнози надає PwC: очікується зростання світового прибутку компаній із приблизно 1 трильйона доларів у 2018 році до майже 16 трильйонів доларів у 2030 році [25].

Окремим напрямом є використання ШІ у кібербезпеці. Ця галузь особливо виграє від впровадження інновацій, адже ШІ допомагає виявляти та аналізувати кіберзагрози, що дозволяє спеціалістам оперативно реагувати на потенційні атаки та підвищувати рівень безпеки. Штучний інтелект аналізує великі обсяги даних, знаходить закономірності, які можуть вказувати на шкідливу активність, і забезпечує виявлення загроз у реальному часі. Водночас компанії мають посилювати заходи безпеки, щоб протистояти ризикам, які можуть зростати

разом із розвитком ШІ. Штучний інтелект можна застосовувати для автоматизації процесів у сфері кібербезпеки, зокрема оновлення та виправлення застосунку, що сприяє забезпеченню захищеності систем. Він також допомагає автоматизувати роботу корпоративних команд реагування на кіберінциденти, підвищуючи їхню ефективність. Використовуючи великі обсяги даних для аналізу та навчання, ШІ здатен пропонувати оптимальні стратегії реагування на різні кіберзагрози. Штучний інтелект у кібербезпеці — це широка сфера, яка дозволяє компаніям не лише зменшувати кіберризики, але й збільшувати доходи завдяки точнішому виявленню загроз та шахрайства. У зв'язку з ускладненням відстеження нових вірусів і шкідливого застосунку, інструменти на основі ШІ стають незамінними для швидкого виявлення та реагування на кіберзагрози. Аналізуючи статистичні дані про кібератаки, ШІ дозволяє визначити найбільш ефективні подальші дії. Штучний інтелект часто виявляється ефективнішим за людину у розпізнаванні шкідливого застосунку. Його інтеграція в багаторівневі системи кібербезпеки, такі як «Інформування про безпеку» та «Управління подіями», допомагає аналітикам покращити моніторинг і виявлення потенційних загроз у корпоративних мережах [25].

Основні напрями використання ШІ для забезпечення кібербезпеки на корпоративному рівні:

- Класифікація даних щодо конфіденційності для дотримання нормативів з їх обробки.
- Профілі кібербезпеки на основі поведінки користувачів.
- Блокування ботів на основі аналізу поведінки користувачів.
- Підвищення готовності до кіберзагроз.
- Виявлення кіберзагроз на основі ШІ.

До потенційних цілей злочинців у кіберпросторі передусім належить мережева інфраструктура бізнес-структур. При цьому «класичні» засоби антивірусної боротьби та кіберзахисту вже не здатні впоратися з такими

епідеміями, і на допомогу приходять рішення на базі штучного інтелекту. Методи ШІ надійні, гнучкі та здатні покращувати системи кіберзахисту на корпоративному рівні від все більшої кількості випереджальних кіберзагроз. Системи штучного інтелекту допомагають посилити кібербезпеку: розпізнають аномалії та нові типи зловмисних програм, сповіщають про кіберзагрози та захищають критичні дані [25].

4.2 Використання штучного інтелекту в інформаційній безпеці держави

Активний розвиток новітніх цифрових технологій, зокрема технологій штучного інтелекту, обумовлює актуальність вивчення проблем інформаційної безпеки, а саме: загроз для інформаційних ресурсів, різних засобів і заходів захисту; бар'єрів для проникнення; вразливих місць у системі захисту інформації. За таких умов особливого значення та актуальності в спектрі суспільних відносин набуває використання штучного інтелекту в системі забезпечення інформаційної безпеки держави в умовах зовнішніх і внутрішніх загроз, оцінювання та аналізу інформаційних загроз і практичне застосування штучного інтелекту в контексті інформаційного опору з метою забезпечення національної безпеки держави, що належить до концептуальних засад суспільства. На сьогодні штучний інтелект займає пріоритетний напрямок у забезпеченні інформаційної безпеки як складової національної безпеки держави.

Застосування технологій штучного інтелекту є одним із чинників, що сприяє забезпеченню інформаційної безпеки держави. Зокрема, штучний інтелект — це ефективний інструмент підвищення кіберстійкості держави, який може бути використано в інформаційній безпеці на державному рівні для захисту комунікаційних, інформаційних і технологічних систем. На даний час застосування штучного інтелекту в інформаційній безпеці значно ширше: аналіз у мережі великого обсягу інформації з метою виявлення та аналітики

інформаційних загроз, прогнозування загроз, передбачення так званих атак нульового дня тощо. Зокрема, штучний інтелект є ефективним помічником у захисті від кіберзагроз. Можна виділити такі напрями використання цифрових продуктів із застосуванням технологій штучного інтелекту з метою забезпечення інформаційної безпеки державних органів, органів місцевого самоврядування, бізнесу:

- виявлення та реагування на кібератаки;
- виявлення шахрайства, зокрема у бізнес-процесах;
- управління подіями безпеки;
- захист кінцевих точок;
- захист додатків і управління вразливостями;
- управління доступом та аутентифікація;
- аналіз поведінки користувачів і пристроїв;
- виявлення шкідливих програм;
- антифішинг [25].

Основні типи технологій штучного інтелекту в системі інформаційної безпеки розглянемо нижче.

4.2.1 Endpoint Detection and Response

EDR (Endpoint Detection and Response) — це платформи, призначені для виявлення атак на кінцевих пристроях, таких як робочі станції, сервери або комп’ютери, а також для швидкого реагування на них. Завдяки штучному інтелекту ці продукти здатні розпізнавати невідомі шкідливі програми, автоматично класифікувати загрози інформаційній безпеці та реагувати на них,

передаючи інформацію в центр управління. Рішення приймаються на основі бази знань, сформованої через збір даних з численних пристроїв.

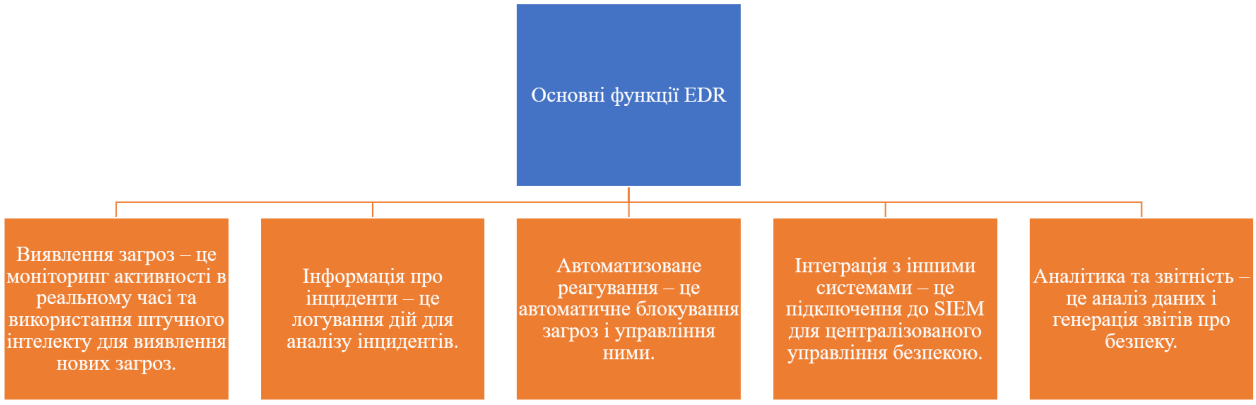

Рисунок 4.2.1 – Основні функції EDR

Крім того, EDR платформи відіграють важливу роль у проактивному управлінні безпекою, включаючи моніторинг активності кінцевих пристроїв в реальному часі. Це дозволяє IT-відділам виявляти аномалії та потенційні вторгнення ще до того, як вони зможуть завдати суттєвої шкоди. Наприклад, за допомогою спеціальних алгоритмів аналітики поведінки, EDR системи здатні виявити нетипові дії користувачів або програм, що може свідчити про спроби зламу.

Ще одним важливим аспектом EDR є їхнє інтеграція з іншими рішеннями з кібербезпеки, такими як SIEM (Security Information and Event Management). Це забезпечує комплексний підхід до безпеки, де дані з різних джерел консолюються для більш точного аналізу та виявлення загроз. Разом ці технології дозволяють організаціям не лише реагувати на інциденти, але й виявляти їхні причини, а також навчатися на основі нових загроз, що постійно виникають у динамічному кіберпросторі. EDR допомагає організаціям виявляти, реагувати й запобігати загрозам, підвищуючи рівень захисту інформаційних активів [26].

4.2.2 Network Detection and Response

NDR (Network Detection and Response) — це платформи й пристрої, що забезпечують виявлення загроз на рівні мережі та оперативне реагування на них. Використовуючи статистику та базу даних про кіберзагрози, вони аналізують мережевий трафік за допомогою ШІ, виявляючи потенційні атаки й автоматично змінюючи конфігурацію мережевих пристроїв і шлюзів. Деякі продукти спеціалізуються на захисті хмарних провайдерів, а також аналізують поштовий трафік, щоб виявити фішингові атаки.

NDR-системи стають ключовим елементом сучасної стратегії кіберзахисту, особливо в умовах постійного зростання кількості загроз і складності атак. Вони забезпечують проактивний підхід до безпеки, аналізуючи величезні обсяги даних у реальному часі. За допомогою ШІ і машинного навчання, NDR не тільки виявляють аномалії, але й класифікують їх за ступенем ризику, дозволяючи адміністраторам швидко приймати рішення щодо реагування.

Особливу роль ці системи відіграють у захисті хмарної інфраструктури, де кількість точок доступу і можливих векторів атак значно більша. Наприклад, NDR-рішення можуть виявляти підозрілу активність, пов'язану з несанкціонованими входами в облікові записи або підозрілими змінами конфігурації. Це важливо для захисту конфіденційних даних та збереження бізнес-операцій [27].

Таблиця 4.2.1 – Приклад порівняння функцій NDR-систем

Функції	Локальні мережі	Хмарна інфраструктура	Аналіз поштового трафіку
Виявлення аномалій	✅	✅	❌
Інтеграція з SIEM	✅	✅	✅

Захист від фішингу	✕	☑	☑
Автоматична зміна конфігурацій	☑	☑	✕

4.2.3 User and Entity Behavior Analytics

UEBA (User and Entity Behavior Analytics) — системи аналізу поведінки користувачів і цифрових сутностей. Вони розпізнають нетипову поведінку, що може свідчити про внутрішні або зовнішні загрози. Технології ШІ у цих системах дозволяють автоматично ідентифікувати аномалії в поведінкових моделях і класифікувати їх як ризики або загрози. Ці системи використовуються для моніторингу доступу, запобігання шахрайству, захисту конфіденційних даних і забезпечення відповідності нормативним вимогам.

Серед ключових застосувань UEBA — моніторинг доступу до критичних систем і даних. Наприклад, система може виявити, якщо користувач раптом почав отримувати доступ до файлів, які раніше не використовував, або здійснює дії поза своїм робочим графіком. Це дозволяє швидко реагувати на загрози, мінімізуючи потенційні збитки. Також UEBA сприяє запобіганню шахрайству, виявляючи нетипову поведінку, пов’язану з фінансовими операціями чи транзакціями.

Інтеграція UEBA із загальною стратегією кіберзахисту підвищує ефективність системи безпеки, забезпечуючи глибокий аналіз і збереження відповідності нормативним вимогам (наприклад, GDPR чи ISO 27001). Завдяки автоматизації аналізу поведінкових даних, ці системи зменшують навантаження на команди безпеки та забезпечують більш точне виявлення прихованих загроз.

4.2.4 Threat Intelligence Platform

TIP (Threat Intelligence Platform) — платформи для раннього виявлення кіберзагроз і реагування на них, які працюють на основі великих обсягів даних і індикаторів компрометації (IoC). Штучний інтелект покращує здатність цих платформ розпізнавати невідомі загрози на ранніх стадіях. Подібно до SIEM-систем, TIP фокусуються на зовнішніх джерелах даних і загрозах.

Рисунок 4.2.2 – Преваги використання Threat Intelligence Platform

4.2.5 Security Information and Event Management

SIEM (Security Information and Event Management) — це рішення для моніторингу інформаційних систем у реальному часі, аналізу подій безпеки з різних джерел і виявлення кіберінцидентів. Застосування ШІ дозволяє ідентифікувати аномалії, зменшувати кількість хибних спрацьовувань і адаптуватися до змін у поведінкових моделях даних. Сучасні кіберзлочинці не атакують безпосередньо IT-інфрастуктуру. Вони діють завуальовано, використовуючи вразливості захисних ресурсів. Такі інциденти залишаються поза увагою, через те, що без «контексту» не вказують на загрозу. Відстежити протиправні дії допомагає постійний моніторинг і аналіз всіх подій, що

відбуваються в IT-інфраструктурі компанії. Таку здатність аналізувати та виявляти інциденти по окремим подіям мають SIEM-рішення.

Наша команда спроєктує і запровадить SIEM систему що дозволить адміністраторам інформаційної безпеки сфокусуватися на реальних загрозах, забезпечуючи їх засобами, що дозволяють оперативно реагувати на загрози безпеки мережі [28].

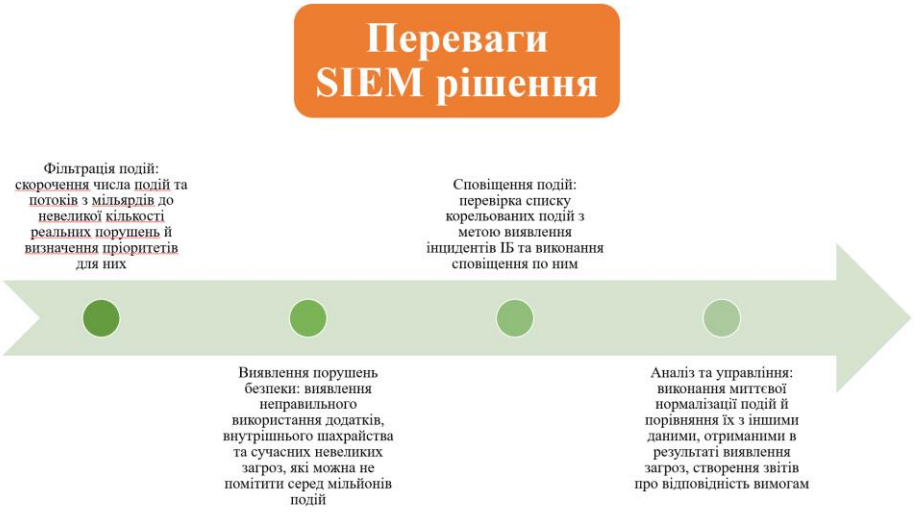

Рисунок 4.2.3 – Переваги SIEM рішень

SIEM (Security Information and Event Management) рішення надає організаціям можливість централізованого збору, аналізу та кореляції даних безпеки з різних джерел у реальному часі, що дозволяє швидко виявляти аномалії та потенційні загрози. Це забезпечує покращене реагування на інциденти, спрощує дотримання регуляторних вимог, дозволяє автоматизувати процеси звітності та аудиту, а також підвищує загальну видимість та контроль над безпекою мережі. В результаті впровадження SIEM рішень організації можуть значно знизити ризики кібератак, покращити управління інцидентами та ефективніше використовувати ресурси безпеки.

4.2.6 Security Orchestration and Automated Response

SOAR (Security Orchestration and Automated Response) — системи, які поєднують виявлення загроз і автоматизацію реагування на кіберінциденти. На відміну від SIEM-систем, SOAR не тільки аналізують події безпеки, але й самостійно виконують дії для нейтралізації загроз.

Рисунок 4.2.4 – Компоненти SOAR

В цілому, SOAR є потужним інструментом для інтеграції інструментів кіберзахисту в єдине рішення, автоматизації процесів захисту та реагування на загрози, а також централізованого управління та аналізу подій, що виникають під час інцидентів безпеки. За допомогою SOAR організації можуть підвищити ефективність своїх заходів кібербезпеки та забезпечити надійніший захист своєї IT-інфраструктури.

4.2.7 Системи захисту застосунків

Системи захисту застосунків (Application Security) — рішення, що виявляють вразливості у прикладних додатках, забезпечують моніторинг та усунення загроз. Завдяки ШІ ці системи автоматично збирають інформацію з відкритих джерел про вразливості та атаки, після чого ініціюють відповідні захисні дії. Згідно з дослідженнями PurpleSec, 2020 року щодня відбувалося 30 тис кібератак. Під час пандемії COVID-19 їхня кількість зросла на 600%. Сучасний бізнес дедалі більше використовує у своїй діяльності програмне забезпечення. Саме воно є вразливим до хакерських нападів. ПЗ керує даними і приймає «розумні» рішення, тож користь бізнесу від нього очевидна. У цьому ж і кіберзлочинці вбачають користь. Тому захист даних — це мастхев для будь-якої сучасної компанії.

Основні компоненти систем захисту додатків включають:

1. Аудит коду та тестування на вразливості. Використання інструментів для аналізу вихідного коду на предмет вразливостей, а також проведення тестування, як, наприклад, пенетраційне тестування.

2. Динамічне тестування застосунків (DAST). Інструменти, які перевіряють працюючі програми на наявність вразливостей в реальному часі, імітуючи атаки з боку зловмисників.

3. Статичне тестування застосунків (SAST). Аналіз вихідного коду або бінарних файлів програми без її виконання для виявлення вразливостей і недоліків.

4. Контроль доступу. Реалізація механізмів аутентифікації та авторизації, які запобігають несанкціонованому доступу до додатків.

5. Захист даних. Шифрування чутливої інформації як в транзиті, так і в стані спочинку, та управління конфіденційністю даних.

6. Системи запобігання втратам (WAF). Веббрандмауери, які моніторять та фільтрують HTTP-трафік між веб-додатками та Інтернетом, захищаючи від поширених веб-загроз, таких як SQL-ін'єкції та атаки типу XSS.

7. Моніторинг та реагування. Збір і аналіз логів та подій безпеки в реальному часі для виявлення та реагування на потенційні загрози.

Впровадження систем захисту додатків є критично важливим для збереження цілісності, конфіденційності та доступності програмного забезпечення, особливо з огляду на зростання складності кібератак та вимоги до регуляторного дотримання.

4.2.8 Антифрод

Антифрод (Antifraud) — системи для виявлення загроз у бізнес-процесах і запобігання шахрайським операціям у режимі реального часу. Використання ШІ дозволяє визначати відхилення від встановлених процесів, що дає змогу оперативно реагувати на можливі фінансові злочини.

Таблиця 4.2.2 – Переваги та недоліки Antifraud

Переваги	**Недоліки**	**Загрози**
Зниження фінансових втрат	Високі витрати на впровадження	Еволюція шахрайських тактик
Захист репутації	Фальшиві позитиви	Некоректні дані для навчання алгоритмів
Аналіз і прогнозування загроз	Складність інтеграції	Кібератаки на антифрод системи
Автоматизація процесів	Технологічні обмеження	Легітимна діяльність, що виглядає підозріло
Дотримання регуляторних вимог	Проблеми з конфіденційністю	Недостатня обізнаність про нові загрози

На ШІ базується технологія діпфейк. Її було використано для створення фейкового відеозвернення Президента України В. Зеленського про капітуляцію, яке було вкинуто в інфопростір у березні 2022 р. З огляду на низьку якість цього «продукту», оперативну реакцію державних комунікацій, Президента України, який особисто спростував фейк, і журналістів, це не мало негативних наслідків щодо громадських настроїв у країні. Відео не досягло своєї мети ні в Україні, ні за кордоном. Загалом ШІ має значний потенціал використання насамперед для створення фото-, аудіо- і відеофейків, а також для роботи ботоферм. ШІ може замінити значну частину персоналу на російських «фабриках тролів», інтернет-бійців, які провокують конфлікти в соцмережах і створюють ілюзію масової підтримки російських наративів користувачами у масштабній війні, розв'язаній проти України. Замість «тролів», що пишуть коментарі, це може робити ШІ з використанням ключових слів і запропонованої йому лексики. Безумовно, визначальний вплив на лояльну аудиторію мають політики, пропагандисти, блогери, конспірологи тощо, а не безіменні боти й інтернет-тролі, але за допомогою ШІ вагу останніх можна штучно збільшити за рахунок кількісного зростання та «тонкого налаштування» під різні цільові аудиторії [25].

ElevenLabs – це американський ШІ-сервіс, який спеціалізується на генерації голосу з високоякісною, людоподібною промовою. Він підтримує 32 мови та ідеально підходить для створення аудіокниг, озвучення відео, реклами та інших форматів. Сервіс дозволяє одночасно слухати та переглядати контент на кількох мовах. По суті ElevenLabs — це онлайн-генератор голосу на базі ШІ, доступ до якого можна отримати прямо з веббраузера. Вже у червні 2023 року компанія залучила $19 млн інвестицій: раунд серії А очолили американський венчурний фонд Andreessen Horowitz, ex-CEO GitHub Нат Фрідман та екскерівник ШІ-відділу в бізнес-інкубаторі Y Combinator Даніель Гросс; вартість компанії тоді оцінили в $100 млн. А вже у січні 2024 року стартапу вдалося отримати $80 млн інвестицій: тоді Bloomberg із посиланням на виконавчого директора проєкту Маті Станішевські повідомив, що оцінка компанії сягнула $1,1

млрд. Так, менше ніж за рік, сервіс ElevenLabs став «єдинорогом». ElevenLabs підходить для різних завдань — від створення аудіокниг і подкастів до озвучування навчальних відео та роботи з віртуальними асистентами. А ще сервіс надає інструменти для налаштування голосу: можна змінювати тон, швидкість та навіть емоції, що робить процес створення голосу гнучким та дозволяє отримати результат, який ідеально збігається з задумом користувача. [29].

Серед ключових функцій ElevenLabs можна виділити наступні:

- висока якість голосу зі штучним інтелектом — сервіс пропонує кілька мовних моделей, які можна вибрати для різних мов і потреб;

- зручне та легке налаштування — в онлайн-інтерфейсі ви можете налаштовувати параметри голосу для всього проєкту: регулювати стабільність, схожість, а також додавати стиль;

- бібліотека голосів — сервіс надає понад 100 готових голосів; також ви можете створити власний голос у розділі VoiceLab; а ще є можливість налаштовувати нові голоси чи завантажувати аудіофайли для клонування;

- простий та інтуїтивний інтерфейс — легко зрозуміти, як все працює; генератор голосу дозволяє змінювати голоси безпосередньо на тій же сторінці, без необхідності використовувати додаткові інструменти.

4.3 Застосування штучного інтелекту у сфері національної безпеки й обороноздатності держави

Сфера оборони та безпеки у світі є провідною галуззю, яка зазнає значного впливу від впровадження технологій штучного інтелекту, що змінюють розстановку сил між державами. Зарубіжний досвід демонструє, що ефективне забезпечення воєнної безпеки та обороноздатності держави в сучасних умовах

залежить від впровадження передових технологій, таких як штучний інтелект і великі дані (Big Data). Важливість застосування ШІ для гарантування національної безпеки підтверджують результати досліджень Науково-технічної організації НАТО. Вони визначають основні технології, що матимуть вирішальний вплив на розвиток воєнної сфери протягом наступних двадцяти років. Серед них ключову роль відіграють Big Data, штучний інтелект, автономні транспортні системи, космічні технології, гіперзвукові апарати, квантові технології, біотехнології, нові матеріали та інші інновації [25].

В Україні наразі ухвалено шість стратегічних програмних документів у сфері безпеки, які безпосередньо або опосередковано стосуються питань національної безпеки, обороноздатності та використання штучного інтелекту, великих даних (Big Data) і сучасних інформаційно-комунікаційних технологій у цих галузях :

Стратегія забезпечення державної безпеки (затверджена Указом Президента України від 16 лютого 2022 р. №56/2022) визначає ключові завдання державної політики у сфері безпеки, серед яких завершення формування та розвиток національної системи кібербезпеки, оптимізація координації її суб'єктів для ефективної протидії кіберзагрозам, а також створення ефективної системи обміну інформацією між суб'єктами безпеки. Вона також передбачає забезпечення доступу до державних інформаційних ресурсів і баз даних [33].

Стратегія національної безпеки України «Безпека людини — безпека країни» (затверджена Указом Президента України від 14 вересня 2020 р. №392/2020) враховує сучасні загрози, зокрема стрімкі технологічні зміни, розвиток штучного інтелекту, квантових, інформаційних і біотехнологій. Основний акцент зроблено на забезпеченні кіберстійкості національної інфраструктури, вдосконаленні систем управління, телекомунікацій, розвідки, логістики та розвитку Збройних Сил на основі сучасних технологій [32].

Стратегія інформаційної безпеки (затверджена Указом Президента України від 28 грудня 2021 р. №685/2021) передбачає протидію дезінформації та інформаційним операціям, створення системи раннього виявлення та прогнозування гібридних загроз, а також розвиток можливостей оборонних структур для захисту інформаційного простору [34].

Стратегія кібербезпеки України «Безпечний кіберпростір — запорука успішного розвитку країни» (затверджена Указом Президента України від 26 серпня 2021 р. №447/2021) визнає забезпечення кібербезпеки пріоритетом національної безпеки. Документ наголошує на необхідності формування ефективної національної системи кіберзахисту, здатної адаптуватися до сучасного цифрового середовища та захищати національний сегмент кіберпростору [35].

Стратегія воєнної безпеки України «Воєнна безпека — всеохоплююча оборона» (затверджена Указом Президента України від 25 березня 2021 р. №121/2021) акцентує увагу на підвищенні боєздатності Збройних Сил, розвитку можливостей кібероборони, а також впровадженні сучасних технологій у сферу озброєння, телекомунікацій та управління військами [36].

Стратегія розвитку оборонно-промислового комплексу України (затверджена Указом Президента України від 20 серпня 2021 р. №372/2021) передбачає створення умов для розвитку оборонно-промислового комплексу через державно-приватне партнерство та міжнародне співробітництво. Пріоритетом є виробництво сучасного озброєння, розвиток технологій штучного інтелекту, нових матеріалів, біотехнологій і систем зв’язку для посилення обороноздатності та експортного потенціалу України [37].

5 АНАЛІЗ АТАК ІЗ ВИКОРИСТАННЯМ ШІ

5.1 Постановка проблеми

Оборона держави, національна безпека та всебічний розвиток суспільства великою мірою залежать від прогресу у сфері високих технологій. Однією з ключових галузей, яка відіграє важливу роль у цьому процесі, є штучний інтелект (ШІ). Більшість людей не мають чіткого уявлення про принципи і алгоритми роботи систем Штучного Інтелекту (СШІ). По суті, ці системи сприймаються як своєрідні «чарівні чорні скриньки», які можуть обробляти природну мову, музику або зображення, адекватно реагуючи на запитання користувачів і надаючи статистично вірні відповіді. При цьому, користувачі, отримуючи результати відповідно до своїх запитів, не усвідомлюють, які джерела лягають в основу відповідей і якими є методи розв'язання завдань. З іншого боку, відсутність прозорих методів перевірки висновків і рекомендацій, запропонованих системами ШІ, викликає невизначеність щодо їх точності та практичної цінності. Це фактично означає, що ШІ може стати елементом інформаційної війни, спрямованої на поширення сумнівних, неперевірених даних та фейків. ШІ може бути потужним інструментом у інформаційних конфліктах, дозволяючи створювати більш переконливі та цілеспрямовані фейкові новини і автоматизувати їх розповсюдження. Варто зазначити, що платформи, які використовують алгоритми рекомендацій на основі штучного інтелекту, застосовуються для визначення пріоритетності контенту з метою маніпуляції емоціями, поглядами та поведінкою користувачів. Штучний інтелект забезпечує ефективну обробку великих обсягів даних (big data) завдяки використанню алгоритмів машинного навчання та аналізу даних. Це дозволяє автоматизувати процеси обробки та аналізу даних, виявляти приховані закономірності, прогнозувати тенденції та патерни, а також оптимізувати процеси прийняття

рішень і створювати інтелектуальні системи управління даними. Отже, системи ШІ сприяють здобуттю знань з великих обсягів даних і дозволяють приймати обґрунтовані рішення на основі їх аналізу [30].

5.2 Основні відомості про вразливості та класифікацію атак на системи штучного інтелекту.

Вразливості систем штучного інтелекту часто пов'язані з характеристиками самих моделей, їхніми даними, архітектурою або інтеграцією в системи. Оскільки ШІ широко використовується в різних сферах, від фінансів до охорони здоров'я, розуміння цих вразливостей є критично важливим для забезпечення безпеки. Нижче наведено порівняльний аналіз основних вразливостей та класифікацій атак (рис. 5.1), реалізація таких аналізів допоможе створити систему зниження ризиків, що дозволить швидше виявляти, готуватися і реагувати на виклики та загрози, які походитимуть від створення та використання систем штучного інтелекту. Атаки на системи штучного інтелекту поділяються на кілька основних типів. Однією з поширених категорій є атаки на етапі навчання (poisoning attacks). У цьому випадку зловмисники модифікують навчальні дані, щоб модель вивчила неправильні закономірності або стала схильною до помилок у певних сценаріях. Інший тип — атаки на етапі виконання (evasion attacks). Зловмисники маніпулюють вхідними даними, щоб обдурити модель, наприклад, змусивши її класифікувати шкідливий об'єкт як безпечний. Також існують атаки, спрямовані на витік інформації (model inversion attacks), де зловмисник намагається відновити приватні дані, використані під час навчання. Ще один вид — атаки на основі перенесення (transfer attacks), де знання про вразливості однієї моделі використовуються для атак на інші, подібні моделі [31].

Категорія	Вразливості	Приклади атак
Дані для навчання	- Нечисті або упереджені дані - Можливість модифікації даних	- Data Poisoning (зміна даних для навчання) - Backdoor Attacks (вбудовування "бекдору" в модель)
Архітектура моделі	- Перенавчання (overfitting) - Чутливість до збурень (adversarial examples)	- Adversarial Attacks (помилкові висновки через зміну входу) - Model Extraction (крадіжка моделі)
Етап використання	- Непередбачуваність результатів - Витік конфіденційної інформації	- Model Inversion (відновлення тренувальних даних) - Membership Inference (визначення участі даних)
Інфраструктура та доступ	- Вразливі API - Обмеження ресурсів	- Denial of Service (DoS, перевантаження системи) - Trojan Attacks (вбудовування шкідливих компонентів)
Конфіденційність	- Збереження чутливих даних у моделі - Відсутність захисту від зворотного аналізу	- Model Inversion - Membership Inference
Доступність	- Високе споживання ресурсів - Низька стійкість до перевантажень	- DoS (Denial of Service) - Атаки на обчислювальну інфраструктуру

Рисунок 5.1 – Порівняльний аналіз

Вразливості систем штучного інтелекту нерозривно пов’язані з конкретними типами атак. Уразливості можуть виникати на різних етапах життєвого циклу системи ШІ — від підготовки даних і навчання моделі до її використання в реальних умовах. Атаки, у свою чергу, спрямовані на експлуатацію цих слабких місць:

- Вразливі дані роблять систему чутливою до атак типу Data Poisoning і Backdoor.

- Чутливість моделі до збурень створює ризики Adversarial Attacks, які можуть суттєво вплинути на її точність.

- Вразливості інфраструктури та доступу можуть використовуватись для атак на доступність, як-от DoS або Trojan Attacks.

Головною стратегією захисту є виявлення та мінімізація вразливостей на всіх рівнях розробки системи. Це включає забезпечення чистоти та цілісності даних, підвищення стійкості моделей до збурень, захист інфраструктури від зовнішніх загроз і використання криптографічних методів для захисту конфіденційної інформації. Таким чином, ефективний захист систем ШІ вимагає комплексного підходу, що враховує як технічні, так і організаційні аспекти [30].

5.3 Огляд атак на системи штучного інтелекту

Першим видом специфічних атак є "Атаки на платформу" (platform attacks), на якій працюють системи штучного інтелекту. Існує три основних різновиди цього типу атак:

- Модифікація даних (data modification) – це маніпулювання параметрами моделі, яке здійснюється шляхом підбору таких вхідних даних, що можуть вивести з ладу внутрішні механізми обробки. Завдяки цим атакам зловмисники впливають на цілісність системи ШІ.

- Відмова в обслуговуванні (denial of service) – це атака, що призводить до виведення системи ШІ з ладу або її уповільнення. Вона реалізується шляхом надсилання великого обсягу трафіку, що заважає доступу звичайних користувачів або уповільнює його. Атаки цього типу безпосередньо впливають на доступність системи.

- Вхідний витік (input leakage) – це захоплення вхідних даних користувачів, що відбувається через компрометацію системи ШІ або експлуатацію вразливостей у її оточенні. Це призводить до порушення конфіденційності даних [31].

Вразливості систем штучного інтелекту тісно пов’язані зі специфічними атаками. Вони можуть виникати на різних етапах життєвого циклу системи ШІ від підготовки даних і навчання моделі до її реального використання. Атаки, у свою чергу, націлені на експлуатацію цих слабких місць:

- Вразливі дані роблять систему чутливою до атак, таких як Data Poisoning і Backdoor.

- Чутливість моделі до збурень створює ризики Adversarial Attacks, які можуть негативно вплинути на її точність.

- Вразливості в інфраструктурі та доступі можуть використовуватись для атак на доступність, як-от DoS або Trojan Attacks.

Отже, ефективний захист систем ШІ потребує комплексного підходу, що враховує як технічні, так і організаційні аспекти.

Другим типом специфічних атак є “Атаки на алгоритм” (algorithm attacks), що використовуються системами штучного інтелекту. Основним їх видом є “Змагальні атаки на ШІ”, які засновані на додаванні “шуму” до вхідних даних, внаслідок чого система робить помилкові передбачення. Один із прикладів цієї атаки ілюструє рис. 5.2, де показані дослідники з бельгійського університету KU Leuven, які зламали відеоаналітичний сервіс із штучним інтелектом, використовуючи кольоровий роздрукований патерн. На зображенні особу зліва система легко розпізнає як людину, тоді як особа праворуч взагалі не ідентифікується як людина, оскільки кольоровий патерн заважає цьому [31].

Рисунок 5.2 – Злам системи відеонагляду

Третім типом специфічних атак є “Атаки на дані” (data attacks). Цей тип атак поділяється на два види. Перший з них - “Атаки з отруєнням даних” (data poisoning attacks). У цих атаках зловмисники експлуатують вразливість, додаючи спеціально створені дані до навчальних наборів систем штучного інтелекту. Хоча ці дані зазвичай не впливають на працездатність системи, їх передача до вхідних даних може призвести до отримання хибного результату, відповідно до очікувань зловмисника. Завдяки отруєнню даних зловмисники порушують цілісність систем штучного інтелекту. Іншим напрямком “Атак на дані” є “Витік даних”.

Атаки, спрямовані на витік навчальних даних, порушують конфіденційність інформації в системах ШІ. Вони намагаються визначити, чи використовувався той чи інший набір даних під час навчання моделі. Існує також ризик, що зловмисник може отримати несанкціонований доступ до особистих даних користувачів, якщо система компрометована і зберігає цю інформацію. Таким чином, здійснюючи атаки витоку даних, кіберзлочинці порушують конфіденційність систем штучного інтелекту [31].

Отже атаки на платформу найбільш руйнівні, але їх найважче реалізувати через інфраструктурні захисти. Атаки на алгоритм дозволяють маніпулювати висновками моделі, проте вимагають високого рівня технічних знань. Атаки на дані є найпростішими, але вони залежать від контролю над джерелами інформації, що ускладнюється відповідними заходами безпеки, наглядно це можна побачити в порівняльній таблиці 5.1

Таблиця 5.1 – Порівняльний аналіз атак на ШІ

Категорія атак	Переваги атак для зловмисника	Недоліки атак (обмеження)
Атаки на платформу	- Дозволяють отримати повний доступ до системи ШІ. - Можливість змінювати параметри або конфігурацію.	- Необхідний доступ до інфраструктури. - Захищені платформи ускладнюють реалізацію атак.
Атаки на алгоритм	- Можна маніпулювати результатами системи. - Експлуатація слабких місць моделі, наприклад, через Adversarial Attacks.	- Потребує глибокого розуміння архітектури моделі. - Ризик виявлення модифікацій.
Атаки на дані	- Найменше зусиль: достатньо модифікувати дані. - Впливає на якість тренування та висновків.	- Ефективність залежить від контролю над джерелами даних. - Захищені джерела знижують ризики.

Важливо підкреслити необхідність забезпечення безпеки на всіх етапах життєвого циклу розробки систем штучного інтелекту. Щоб не пригнічувати впровадження ШІ, до некритичних систем можна застосовувати менш жорсткі рекомендації та більш гнучкий життєвий цикл, в той час як для критичних систем потрібні строгі й стандартизовані процедури. Варто акцентувати увагу на важливості прозорості, тестування, підзвітності алгоритмів та осіб, які їх розробляють і надають послуги з використанням ШІ. У разі відмов критичних систем штучного інтелекту розробники та оператори повинні нести відповідальність за завдану шкоду. Але ця рекомендація потребує більш детальної розробки з огляду на потенційно серйозні наслідки. Безсумнівно, системи ШІ створюють нові цінності для суспільства. Для подальшого розвитку вкрай важливо забезпечити безпечність технологій штучного інтелекту як під час їх розробки, так і в процесі експлуатації.

ВИСНОВКИ

У процесі дослідження теми інформаційних ризиків використання штучного інтелекту в інфокомунікаційних мережах було проаналізовано сучасні тенденції, виклики та можливості, пов’язані з інтеграцією технологій штучного інтелекту (ШІ) в інформаційні та комунікаційні системи. Робота спрямована на виявлення загроз, які виникають унаслідок впровадження ШІ, а також на розробку рекомендацій для їх мінімізації. Штучний інтелект, як ключова технологія сучасності, одночасно створює нові можливості для розвитку інфокомунікаційних мереж і породжує значні ризики. Проведена оцінка та порівняльний аналіз загроз допомогли структурувати інформацію про потенційні небезпеки, пов'язані з використанням ШІ, та визначити пріоритетні напрями для роботи над їх мінімізацією. Важливим є розуміння подвійної природи ШІ: як загрози, так і засобу для забезпечення безпеки. ШІ може стати потужним інструментом для зміцнення кіберзахисту, але це потребує комплексного підходу до його впровадження, включно з етичними, правовими та технічними аспектами. Основними результатами дослідження є:

- Ідентифікація ризиків використання ШІ. У ході роботи було визначено ключові ризики, включно з питаннями кібербезпеки, конфіденційності даних, надмірної залежності від технологій, а також можливості використання ШІ в якості інструменту для реалізації кібератак.

- Оцінка та порівняльний аналіз загроз ШІ. Було проведено оцінку загроз, що виникають при впровадженні ШІ, із застосуванням кількісних та якісних методів. Порівняльний аналіз виявив, що основні ризики варіюються залежно від сценарію використання у технологічних системах загрози пов'язані з технічними помилками та експлуатацією вразливостей та соціальних системах ризики включають маніпуляцію даними, дезінформацію та порушення етичних

принципів. Цей аналіз дозволив систематизувати інформацію про загрози та виділити найбільш критичні аспекти для розробки методів їх мінімізації.

- Штучний інтелект як цифровий інструментарій забезпечення безпеки функціонування. У роботі також підкреслено, що ШІ може використовуватись як ефективний інструмент для підвищення безпеки. Перша це системи виявлення загроз, алгоритми ШІ здатні аналізувати великі обсяги даних у реальному часі для виявлення аномалій і кібератак. Друге - прогнозування ризиків за допомогою ШІ можливо передбачати потенційні уразливості та впроваджувати превентивні заходи. Та автоматизація процесів захисту – це інтеграція інтелектуальних рішень у системи безпеки дозволяє значно скоротити час реагування на загрози.

Незважаючи на ці переваги, важливо забезпечити прозорість і надійність систем, керованих ШІ.

- Загрози та використання ШІ. Було окреслено три ключові аспекти ризиків, пов’язаних із ШІ.
- Загрози від зловмисного використання ШІ: застосування технологій для створення шкідливих програм, дезінформації або автоматизації атак.
- Непередбачуваність ШІ: складність прогнозування результатів дій ШІ через складність моделей, що використовуються.
- Соціальні ризики: зниження довіри до систем через потенційні порушення прав користувачів і зловживання даними. Також було розглянуто позитивні аспекти використання ШІ, які сприяють розвитку захищеності інфокомунікаційних мереж.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Штучний інтелект: проблеми та перспективи правового регулювання в Україні та ЄС [Електронний ресурс] // «Юридична практика». – 2023. – Режим доступу до ресурсу: https://pravo.ua/shtuchnyi-intelekt-problemy-ta-perspektyvy-pravovoho- rehuliuvannia-v-ukraini-ta-ies/.

2. Штучний інтелект та технологія GPT: юридичні тонкощі для користувачів [Електронний ресурс] // Анастасія Клян, старша юристка судової практики GOLAW. – 2024. – Режим доступу до ресурсу: https://jurliga.ligazakon.net/news/226223_shtuchniy-ntelekt-ta-tekhnologya-gpt-yuridichn-tonkoshch-dlya-koristuvachv.

3. CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE No 2 (22) 2023ISSN 2663-4023. // Кібербезпека: освіта, наука, техніка / – Київ: Київський університет імені Бориса Грінченка, 2023. – (№2).

4. Тренди ШІ: які етичні загрози несе використання штучного інтелекту [Електронний ресурс] // Юрій Гайдай Старший економіст CES. – 2023. – Режим доступу до ресурсу: https://speka.media/trendi-si-yaki-eticni-zagrozi-nese-vikoristannya-stucnogo-intelektu-v4q3wp.

5. Штучний інтелект: три головні загрози [Електронний ресурс] // DOU. – 2023. – Режим доступу до ресурсу: https://dou.ua/forums/topic/46254/.

6. Why We Need to See Inside AI’s Black Box [Електронний ресурс] // Saurabh Bagchi & The Conversation US. – 2023. – Режим доступу до ресурсу: https://www.scientificamerican.com/article/why-we-need-to-see-inside-ais-black-box/.

7. Штучний інтелект в юриспруденції: очікується стрімке зростання [Електронний ресурс] // Петро Білик, керівник практики Технологій та інвестицій Juscutum. – 2023. – Режим доступу до ресурсу:

https://www.juscutum.com/news/shtuchniy-intelekt-v-yurisprudenciyi-%20ochikuietsya-strimke-zrostannya.

8. Штучний інтелект: загрози і можливості [Електронний ресурс] // EPAM. – 2023. – Режим доступу до ресурсу: https://careers.epam.ua/blog/artificial-intelligence-threats-and-opportunities.

9. РОЗПОРЯДЖЕННЯ, Про схвалення Концепції розвитку штучного інтелекту в Україні [Електронний ресурс] // КАБІНЕТ МІНІСТРІВ УКРАЇНИ. – 2020. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%80#Text.

10. Регулювання штучного інтелекту в Україні: Мінцифри презентувало дорожню карту [Електронний ресурс] // Міністерство цифрової трансформації України. – 2023. – Режим доступу до ресурсу: https://www.kmu.gov.ua/news/rehuliuvannia-shtuchnoho-intelektu-v-ukraini-mintsyfry-prezentuvalo-dorozhniu-kartu.

11. Італія першою з країн заходу заблокувала ChatGPT [Електронний ресурс] // Суспільне новини, Надія Собенко. – 2023. – Режим доступу до ресурсу: https://suspilne.media/432498-italia-persou-z-krain-zahodu-zablokuvala-chatgpt/.

12. ЗАКОН УКРАЇНИ Про захист персональних даних [Електронний ресурс] // Верховна Рада України. – 2024. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2297-17#Text.

13. ЗАКОН УКРАЇНИ Про інформацію [Електронний ресурс] // Верховна Рада України. – 2023. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2657-12#Text.

14. AI Act [Електронний ресурс] // Digital Strategy. – 2024. – Режим доступу до ресурсу: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

15. Захист персональних даних у США [Електронний ресурс] // Сергій Флорескул, Віолетта Лосєва. – 2024. – Режим доступу до ресурсу: https://www.avitar.legal/post/zahist-personalnih-danih-u-ssha .

16. Захист даних у Канаді та США: чого чекати у 2023. CCPA, CRPA, PIPEDA та BILL 27-C [Електронний ресурс] // Юлія Пустовіт IT юристка в Legal IT Group. – 2023. – Режим доступу до ресурсу: https://legalitgroup.com/zahist-danih-u-kanadi-ta-ssha-chogo-chekati-u-2023-ccpa-crpa-pipeda-ta-bill-27-c/?gad_source=1&gclid=CjwKCAiAxea5BhBeEiwAh4t5Kw99H5oHMIKvanwpP44RKoC6xf1Lr1odJPcEpF56Z_Rfbs36ampJHRoCyjQQAvD_BwE.

17. Personal Information Protection and Electronic Documents Act [Електронний ресурс] // Published by the Minister of Justice. – 2024. – Режим доступу до ресурсу: https://laws-lois.justice.gc.ca/eng/acts/P-8.6/.

18. Закон України «Про електронні довірчі послуги» [Електронний ресурс] // Верховна Рада України.. – 2017. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2155-19#Text.

19. Використання ChatGPT: Які потенційні ризики? [Електронний ресурс] // Крістіан Перрі. – 2023. – Режим доступу до ресурсу: https://undetectable.ai/blog/uk/%D0%B2%D0%B8%D0%BA%D0%BE%D1%80%D0%B8%D1%81%D1%82%D0%B0%D0%BD%D0%BD%D1%8F-chatgpt-%D1%8F%D0%BA%D1%96-%D0%BF%D0%BE%D1%82%D0%B5%D0%BD%D1%86%D1%96%D0%B9%D0%BD%D1%96-%D1%80%D0%B8%D0%B7%D0%B8%D0%BA%D0%B8/.

20. Як виявити штучний інтелект і зробити його непомітним [Електронний ресурс] // Крістіан Перрі. – 2023. – Режим доступу до ресурсу: https://undetectable.ai/blog/uk/%D1%8F%D0%BA-%D0%B2%D0%B8%D1%8F%D0%B2%D0%B8%D1%82%D0%B8-%D0%B0%D0%B9-%D0%BF%D0%B8%D1%81%D1%8C%D0%BC%D0%BE/.

21. Як захиститися від фішингових атак і повідомляти про них [Електронний ресурс] // Пошук Google Довідка. – 2024. – Режим доступу до ресурсу: https://support.google.com/websearch/answer/106318?hl=uk.

22. Проблеми забезпечення безпеки в комп'ютерних системах і мережах. Типова корпоративна мережа. Засоби захисту мереж [Електронний ресурс] // Оксана Десятник. – 2020. – Режим доступу до ресурсу: https://classmill.com/659/112/m/nr18Q.

23. Штучний інтелект (ШІ) – що це таке, як працює і навіщо потрібен. [Електронний ресурс] // Termin. – 2023. – Режим доступу до ресурсу: https://termin.in.ua/shtuchnyy-intelekt/.

24. Роль штучного інтелекту в кібербезпеці [Електронний ресурс] // Андрій Боренков Партнер, керівник Advisory. – 2024. – Режим доступу до ресурсу: https://www.bdo.ua/uk-ua/insights-2/information-materials/2024/the-role-of-ai-in-cybersecurity-anticipating-and-preventing-attacks.

25. ШТУЧНИЙ ІНТЕЛЕКТ І БЕЗПЕКА – Київ: ТОВ «Консалтингова компанія «СІДКОН», 2024. – 295 с. – (КОГУТ Ю.І.). – (1).

26. Захист кінцевих точок: як не загубитися у різноманітті продуктів [Електронний ресурс] // Trellix. – 2023. – Режим доступу до ресурсу: https://trellix.bakotech.com/home/endpoint-protection-how-not-to-get-lost-in-the-variety-of-products.

27. Network Detection and Response (NDR) [Електронний ресурс] // Vectra. – 2022. – Режим доступу до ресурсу: https://www.vectra.ai/topics/network-detection-and-response.

28. SIEM Системи управління інформаційною безпекою та подіями інформаційної безпеки [Електронний ресурс] // Softlist. – 2023. – Режим доступу до ресурсу: https://softlist.ua/servises/siem.

29. ШІ-сервіс ElevenLabs переклав інтерв'ю Зеленського Лексу Фрідману. Що це за стартап та як він працює? [Електронний ресурс] // Дмитро Казанцев. – 2025. – Режим доступу до ресурсу: https://mezha.media/articles/shi-servis-elevenlabs-iak-vin-pratsiuie/.

30. ЗАГРОЗИ ТА РИЗИКИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ [Електронний ресурс] // Національна академія Служби безпеки України. – 2023. – Режим доступу до ресурсу: https://drive.google.com/file/d/1hBhSmuti_7G7HhRC_BC60sl8kv3yJrdf/view?usp=sharingq1eX3tIhl_N-5oOi_14AEr0JhH0zA4lhA.

31. ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ: АНАЛІЗ ВРАЗЛИВОСТЕЙ, АТАК І КОНТРЗАХОДІВ [Електронний ресурс] // Національний аерокосмічний університет ім. М. Є. Жуковського. – 2022. – Режим доступу до ресурсу: https://drive.google.com/file/d/1vKpGTnCID0Wujtge0mXoRM-loldKRJLg/view?usp=sharing.

32. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №392/2020 Про рішення Ради національної безпеки і оборони України від 14 вересня 2020 року «Про Стратегію національної безпеки України» [Електронний ресурс] // Рада національної безпеки і оборони України. – 2020. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/3922020-35037.

33. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №56/2022 Про рішення Ради національної безпеки і оборони України від 30 грудня 2021 року «Про Стратегію забезпечення державної безпеки» [Електронний ресурс] // Рада національної безпеки і оборони України. – 2022. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/562022-41377.

34. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №685/2021 Про рішення Ради національної безпеки і оборони України від 15 жовтня 2021 року \"Про Стратегію інформаційної безпеки\" [Електронний ресурс] // Рада національної безпеки і

оборони України. – 2021. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/6852021-41069.

35. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №447/2021 Про рішення Ради національної безпеки і оборони України від 14 травня 2021 року \"Про Стратегію кібербезпеки України\" [Електронний ресурс] // Рада національної безпеки і оборони України. – 2021. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/4472021-40013.

36. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №121/2021 Про рішення Ради національної безпеки і оборони України від 25 березня 2021 року «Про Стратегію воєнної безпеки України» [Електронний ресурс] // Рада національної безпеки і оборони України. – 2021. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/1212021-37661.

37. УКАЗ ПРЕЗИДЕНТА УКРАЇНИ №372/2021 Про рішення Ради національної безпеки і оборони України від 18 червня 2021 року «Про Стратегію розвитку оборонно-промислового комплексу України» [Електронний ресурс] // Рада національної безпеки і оборони України. – 2021. – Режим доступу до ресурсу: https://www.president.gov.ua/documents/3722021-39733.

38. Як виявити ChatGPT: Інструменти та поради для виявлення [Електронний ресурс] // Крістіан Перрі. – 2023. – Режим доступу до ресурсу: https://undetectable.ai/blog/uk/%D1%8F%D0%BA-%D0%B2%D0%B8%D1%8F%D0%B2%D0%B8%D1%82%D0%B8-chatgpt/.