

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІВЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЮРІЯ ФЕДЬКОВИЧА**

**Факультет історії, політології та міжнародних відносин
Кафедра міжнародних відносин та суспільних комунікацій**

**ПРОБЛЕМА РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ: ВИКЛИКИ
ТА МОЖЛИВІ РІШЕННЯ**

Кваліфікаційна робота

Рівень вищої освіти – другий (магістерський)

Виконав:

студент 6 курсу, 604 групи

Щур Орест Михайлович

Керівник:

кандидатка історичних наук,

доцентка **Звоздецька О.Я.**

Рецензент:

кандидат політичних наук,

доцент **Швидюк С.М.**

До захисту допущено

на засіданні кафедри

протокол №_____ від _____ 2024 р.

Зав. кафедрою _____ проф. Макар В.Ю.

Чернівці – 2024

Анотація

Кваліфікаційна робота присвячена аналізу проблем регулювання Штучного інтелекту (ШІ) у контексті викликів, що виникають у сфері прав людини, кібербезпеки та демократії. Дослідження охоплює міжнародні та національні підходи до регулювання ШІ, зокрема ініціативи ЄС, США, Великобританії та України та наголошує на інтернаціональній співпраці. Особливу увагу приділено ризикам використання ШІ, таким як маніпуляція даними, загрози конфіденційності та створення дискримінаційних алгоритмів. Розглянуто також потенціал ШІ для запобігання кіберзлочинності та підвищення ефективності рішень у ключових сферах.

Ключові слова: Штучний інтелект, регулювання, права людини, кібербезпека, конфіденційність, міжнародне право, автоматизація, інновації.

Summary

The qualification work is devoted to the analysis of the problems of regulating Artificial Intelligence (AI) in the context of challenges arising in the field of human rights, cybersecurity and democracy. The study covers international and national approaches to regulating AI, in particular initiatives of the EU, the USA, the UK and Ukraine, and emphasizes international cooperation. Particular attention is paid to the risks of using AI, such as data manipulation, privacy threats and the creation of discriminatory algorithms. The potential of AI for preventing cybercrime and increasing the efficiency of decisions in key areas is also considered.

Keywords: Artificial intelligence, regulation, human rights, cybersecurity, privacy, international law, automation, innovation.

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів наукових досліджень інших авторів мають посилання на відповідне джерело.

(підпис) Щур О.М.

ЗМІСТ

<u>ВСТУП</u>	4
<u>РОЗДІЛ 1 КОНЦЕПТУАЛЬНО-ТЕОРЕТИЧНІ АСПЕКТИ ДОСЛІДЖЕННЯ</u>	7
<u>1.1. Підходи та розуміння поняття «Штучного інтелекту»</u>	7
<u>1.2. Стан джерельної бази дослідження</u>	19
<u>РОЗДІЛ 2 ВИКЛИКИ ТА НАСЛІДКИ ВИКОРИСТАННЯ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ</u>	28
<u>2.1. Штучний інтелект та кібербезпека: ризики та можливості</u>	28
<u>2.2. Вплив Штучного інтелекту на права людини, демократію та верховенство права</u>	40
<u>2.3. Вплив ШІ на захист інтелектуальної власності та авторських прав</u>	50
<u>РОЗДІЛ 3 ПРАВОВІ АСПЕКТИ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ</u>	62
<u>3.1. Тенденції розвитку зарубіжного та вітчизняного законодавства щодо регулювання систем Штучного інтелекту</u>	62
<u>3.2. Інструменти та механізми міжнародного -правового регулювання систем Штучного інтелекту</u>	80
<u>3.3. Глобальна ініціатива з управління Штучним інтелектом</u>	94
<u>ВИСНОВКИ</u>	100
<u>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</u>	104
<u>ДОДАТКИ</u>	121
<u>SUMMARY</u>	125

ВСТУП

Актуальність дослідження. З розвитком інформаційних та комунікаційних технологій (ІКТ) сучасний світ зазнав значних змін, які торкнулися майже всіх сфер життя. Розробка швидкісного інтернету, таких як 4G та 5G, відкрила нові можливості для передачі даних та створення глобальних мереж. Завдяки цьому цифровізація стала основою для трансформації багатьох галузей, від освіти до медицини, від бізнесу до культури. Високошвидкісний Інтернет створює взаємодію людей і пристроїв у реальному часі, сприяючи розвитку концепції «Інтернету речей», яка забезпечує інтеграцію побутових, виробничих і навіть транспортних систем.

Вадливим аспектом цифровізації є використання хмарної технології. Хмарні сховища стали не лише місцем для збереження даних, а й інструментом для роботи з ними у віддаленому доступі. Бізнес активно використовує хмарні обчислення для оптимізації витрат і зростання ефективності. Освітні заклади, державні установи та приватні підприємства можуть інтегрувати хмарні сервіси для автоматизації процесів, спрощення документообігу та зменшення фізичного використання серверів. Це дозволяє значно економити ресурси та забезпечувати доступність даних будь-де і будь-коли.

Однак найвагомішим кроком у цьому технологічному поступі є впровадження Штучного інтелекту (ШІ). Завдяки вмінню аналізувати величезні обсяги інформації та створювати інтелектуальні рішення, ШІ став важливим інструментом у медицині, прогнозуванні клімату, автоматизації виробництва та навіть у творчих галузях. Його можливості сприяють підвищенню точності та швидкості обробки інформації, що сприяє прийняттю ефективних рішень. Наприклад, медичні системи на основі ШІ допомагають діагностувати хвороби на ранніх стадіях, тоді як автоматизовані програми в бізнесі дозволяють оптимізувати операції та підвищувати прибутковість.

Сьогодні розвиток ІКТ є двигуном глобальних змін. Технології 5G, хмарні платформи та ШІ створюють нову реальність, у якій межі між фізичним

і цифровим світом стають дедалі розмитішими. Цей прогрес відкриває перед людством унікальні можливості, але водночас ставить питання про кібербезпеку, етику використання ШІ та збереження приватності. Важливо, щоб цей технологічний стрибок був спрямований на благо суспільства та забезпечував стає майбутнє для всіх.

Тема регулювання Штучного інтелекту (ШІ) на сьогоднішній день є вкрай актуальною і важливою з кількох ключових причин.

По-перше, ШІ швидко зростає і впроваджується у різноманітні сфери, від виробництва до медицини, від освіти до фінансів. Це відкриває безліч можливостей для покращення ефективності, інновацій та розвитку економіки. Проте одночасно з цим зростають прямі ризики, що тісно переплітаються з безпекою, прозорістю та етичним використанням ШІ.

По-друге, відсутність адекватного регулювання призводить вкрай важливих проблем, таких як порушення приватності, дискримінація або навіть загрози для національної безпеки. Законодавство та політика в галузі ШІ поки що розвиваються дуже повільно, що створює потребу у глибокому аналізі і розробці ефективних регуляторних механізмів.

По-третє, зростаюча усвідомленість суспільства щодо впливу ШІ та його потенційних наслідків також підсилює попит на належне регулювання. Громадські організації, академічні установи і галузеві лідери активно висловлюють свої погляди щодо необхідності створення балансу між новими технологіями та захистом прав і свобод людини.

Отже, дослідження на тему регулювання ШІ не лише вирішує актуальні проблеми сьогодення, але й має значний потенціал для внесення вагомому внеску в розвиток наукових знань та формування практичних рекомендацій для політичних та законодавчих органів.

Метою дослідження є аналіз проблеми регулювання Штучного інтелекту, окреслення його викликів і перспектив.

Відповідно до визначеної мети у роботі поставлено **такі завдання**, як:

- дослідити підходи та розуміння поняття «Штучного інтелекту»;
- проаналізувати виклики і загрози, що породжує використання Штучного інтелекту;
- охарактеризувати вплив Штучного інтелекту на права людини, демократію та верховенство права;
- окреслити інструменти та механізми міжнародного-правового регулювання функціонування систем ШІ
- визначити тенденції розвитку зарубіжного та вітчизняного законодавства щодо регулювання систем ШІ.

Об'єкт дослідження – Штучний інтелект, як складова процесу цифровізації.

Предмет дослідження – проблема правого регулювання Штучного інтелекту.

Методи дослідження: Для розкриття предмету дослідження магістерської роботи автором використано певні загальнонаукові та спеціальні методи, зокрема: порівняльний та історичний методи, метод аналізу, індукції та дедукції, метод узагальнення.

Структура роботи – магістерська робота виконана на 132 сторінках (з них основного тексту – 100 сторінок), складається зі вступу, 3-х розділів, висновків, списку використаних джерел на 16 сторінках (125 джерел).

РОЗДІЛ 1 КОНЦЕПТУАЛЬНО-ТЕОРЕТИЧНІ АСПЕКТИ ДОСЛІДЖЕННЯ

1.1. Підходи та розуміння поняття «Штучного інтелекту»

Активне використання Штучного інтелекту (ШІ) змінює економіку, обіцяючи підвищити продуктивність, ефективність та знизити витрати. ШІ підвищує рівень життя, допомагає людям робити кращі прогнози та приймати ґрунтовні рішення. Однак ці технології все ще перебувають у зародковому стані, і для Штучного інтелекту залишається багато перспектив для вирішення глобальних проблем і сприяння інноваціям. У той же час технології ШІ викликають занепокоєння та породжують етичні проблеми. Для розуміння проблеми дослідження необхідно простежити історію зародження та розвитку ШІ, а також підходи до розуміння цього поняття.

Штучний інтелект (ШІ) є предметом багатьох досліджень вітчизняних і зарубіжних вчених. Однак поняття ШІ все ще залишається неоднозначним у сучасній науці. Вперше термін «Штучний інтелект» був використаний на семінарі в Стенфордському університеті в 1956 р., але ШІ є сферою досліджень з 1940-х рр. [57].

Термін «Штучний інтелект» має як широке, так і вузьке значення. У вузькому сенсі ШІ – «це інформація, що генерується автоматизованими системами або функціями програм комп'ютера». В широкому – «Штучний інтелект - це здатність автоматизованих систем або програм комп'ютера виконувати людські функції на основі аналізу зовнішніх факторів і створювати конкретні рішення, спираючись на життєвий досвід людини [3, с. 2].

Існують два наукові методи створення ШІ, які мають на меті різні завдання. У першому випадку головним завданням вченого є розробка системи, подібної до людського мислення. Представники другого методу зосереджуються переважно на вирішенні завдань автоматизованих систем і

програм комп'ютера. Для них не існує суттєвої абсолютної відповідності між алгоритмами системи та людським мисленням [57].

Дослідження ШІ в сучасному розумінні почалися незабаром після появи перших комп'ютерів; з кінця 50-х рр. XX ст. дослідники Штучного інтелекту намагалися розробити розумні машини, що імітують роботу мозку.

Історія ШІ почалася в 1950 р., коли була створена перша символна мова для оброблення переліків - мова програмування IPL1. Незабаром після цього була написана перша програма, що призвела до досягнень у галузі ШІ. Програма отримала назву Logic Theorist (1956), що призначалася для механізованого доведення теорем в обчислювальних твердженнях.

Алгоритм управління намагається зменшити різницю між оцінкою поточної ситуації та оцінкою однієї з цілей або підцілей. Своєю чергою, цей метод часто використовувався як теоретична основа для шахових програм: пошук заданої глибини ходу, функції оцінки позиції та обмеження області пошуку. На цих принципах і сьогодні базуються всі шахові програми.

Перша шахова програма була завершена в 1957 р. Її поведінка ґрунтується на евристиках (правилах, які можна вибрати без точної теоретичної бази). У 1966 р. була представлена програма для імітації людської мови, так звана програма віртуального діалогу. A.L.I.C.E. (1995) - найновіша діалогова програма, яка може спілкуватися природною мовою. Однак вона так і не пройшла тест Тюрінга [57].

Тест Тюрінга — це експеримент, запропонований британським математиком і криптографом Аланом Тюрінгом у 1950 році, щоб визначити, чи може машина думати, як людина. Його ідея була викладена в статті «Computing Machinery and Intelligence». Суть тесту полягає в тому, щоб оцінити, чи здатна комп'ютерна система вести діалог таким чином, що людина-експериментатор не зможе відрізнити її відповіді від відповідей іншої людини [57].

В 1966 р. комп'ютерний науковець Масачусетського технологічного інституту Джозеф Вайзенбаум створив ELIZA, що вважається першим чат-

ботом призначеним для імітації терапії шляхом перепрофілювання відповідей користувачів на питання, які спонукали до подальшої розмови — також відомий як Аргумент Роджеріана.

Аргумент Роджеріана — це метод переконання, заснований на ідеях психолога Карла Роджерса. Він використовується для вирішення конфліктів або обговорення спірних питань. Головна мета цього підходу — не стільки переконати опонента, скільки досягти взаєморозуміння, поваги та компромісу [46].

Між 1966 і 1972 рр. Центр Штучного інтелекту при Стенфордській дослідницькій ініціативі (The Stanford Artificial Intelligence Lab) розробив робота Shakey — мобільного робота, оснащеного датчиками та телекамерою, який використовували для навігації в різних середовищах. Мета створення Shakey полягала в «розробці концепцій і методів ШІ - автомату функціонувати незалежно в реалістичних середовищах», згідно з документом, опублікованим пізніше.

Документ «SHAKEY: Experimentation in Robot Learning and Planning» (Експерименти в навчанні та плануванні роботів) був опублікований у 1969 році в рамках досліджень Штучного Інтелекту, які проводилися в SRI International. Мета проєкту формулювалася як "розробка концепцій і методів Штучного інтелекту для забезпечення автономної роботи автоматів у реалістичних середовищах". У цьому документі описано створення алгоритмів, сенсорних систем та інтеграцію планування й виконання завдань для досягнення автономності Shakey [96].

Після Дартмутської конференції в 1950-х рр. дослідження ШІ почали розвиватися в таких поважних установах, як Массачусетський технологічний інститут (MIT), Стенфордський університет (Stanford) і Університет Карнегі (Carnegie Mellon). Інструментальні особи, які стояли за цією роботою, потребували нагод для обміну інформацією, ідеями та відкриттями. З цією

метою в 1977 р. та в 1979 р. було проведено Міжнародну сп конференцію з ШІ, але більш об'єднана організація ще не виникла.

Американська асоціація ШІ була створена в 1980-х рр., для заповнення цієї прогалини. Організація зосередилася на створенні журналу в цій галузі, проведенні семінарів і плануванні щорічної конференції. Товариство перетворилося на Асоціацію з розвитку Штучного інтелекту (Association for the Advancement of Artificial Intelligence AAAI) є некомерційною організацією в сфері науки, яка займається просуванням наукового розуміння механізмів, що лежать в основі мислення та інтелектуальної поведінки, їх втілення в машинному механізмі [108].

У 1974 р. прикладний математик сер Джеймс Лайтхілл опублікував критичну доповідь про академічні дослідження ШІ, вважаючи, що дослідники, по суті, перевищили й занизили обіцянки, коли мова розпочалася про потенційний інтелект машин. Його засудження призвело до різкого скорочення фінансування.

Період між кінцем 1970-х і поч. 1990-х ознаменував «зиму Штучного інтелекту» (термін, вперше використаний у 1984 році), який означав розрив між очікуваннями ШІ та недоліками технології. Лише наприкінці 1990-х ця сфера отримала більше фінансування на дослідження та розробки, щоб зробити суттєві стрибки вперед [105]. Зокрема, в 1996 р. комп'ютерна система IBM Deep Blue - програма для шахової гри - змагалася з тогочасним чемпіоном світу з шахів Гарі Каспаровим у матчі з шести партій [104].

Однак, значного зростання, саме з 2000 р. коли у Лабораторії ШІ Массачусетського технологічного інституту під керівництвом доктора Синтії Брізіл було створено робота Kismet, що мав датчики, мікрофон і програмування, яке окреслює «процеси людських емоцій». Усе це допомогло роботі зчитувати та імітувати різноманітні почуття.

Після того, як програма Deep Blue від IBM успішно перемогла чемпіона світу з шахів, у 2011 році компанія створила ще одну конкурентоспроможну

комп'ютерну систему, яка продовжила грати в популярну американську вікторину Jeopardy [121 с. 2].

Важливий технічний прорив у розвитку ШІ було здійснено науковцем Джеффрі Гінтона (Geoffrey Hinton) в 2012 р., який почав досліджувати ідею нейронних мереж (системи ШІ, створеної для обробки даних подібно до людського мозку). Робота Гінтона над нейронними мережами та глибоким навчанням — процесом, за допомогою якого система ШІ вчиться обробляти велику кількість інформації тоздійснювати точне прогнозування — стала основою для таких процесів ШІ, як обробка природних мов та розпізнавання мовлення [105]. Ажіотаж навколо роботи Гінтона призвів до того, що він приєднався до Google у 2013 р. Однак у 2023 р. пішов у відставку, щоб мати змогу вільніше говорити про небезпеку створення ШІ.

Робототехніка зробила значний стрибок уперед із перших днів створення Kismet, коли гонконгська компанія Hanson Robotics у 2016 р. створила Софію, «людиноподібного робота», здатного до міміки, жартів та розмов. Завдяки її інноваційному Штучному інтелекту та здібностям спілкуватися з людьми, Софія стала світовим феноменом і регулярно з'являлася в ток-шоу. Ситуацію ускладнювало і те, що в 2017 р. Саудівська Аравія надала Софії громадянство, що зробило її першою істотою зі ШІ, яка отримала це право. Цей крок викликав значну критику серед жінок Саудівської Аравії, які не мали певних прав, які мала Софія [105].

Сплеск інтересу та вдосконалення ШІ в останні роки значною мірою відбувся завдяки розробкам генеративного ШІ — або здатності створювати текст, зображення та відео у відповідь на текстові підказки [114 с.8].

Важливе місце в роботі займає дослідницька компанія ШІ OpenAI, яка заснована в грудні 2015 р. Ілоном Маском, Семом Альтманом, Іллею Суцкевером та іншими для створення безпечного ШІ на користь людству. Спочатку заснована як некомерційна організація, тепер має гібридну структуру, де OpenAI LP функціонує під наглядом некомерційної OpenAI Inc.. Компанія

часто співпрацює з Microsoft, яка є її головним партнером. Штаб-квартира OpenAI розташована в Сан-Франциско, США [88].

Дослідницька компанія OpenAI створила генеративний попередньо навчений трансформатор (GPT), який став архітектурною основою для її ранніх мовних моделей GPT-1 і GPT-2, які навчалися на мільярдах вхідних даних. Навіть з такою кількістю навчання їх здатність генерувати характерні текстові відповіді була обмежена. Натомість GPT-3 викликала дедалі більше галасу, коли вона була випущена в 2020 р., і стала сигналом про серйозний розвиток ШІ. GPT-3 було навчено за 175 мільярдами параметрів, що значно перевищило 1,5 млрд. параметрів, за якими було навчено GPT-2.

2023 р. став знаменним роком у історії генеративного ШІ. OpenAI не лише випустив GPT-4, який знову ґрунтується на потужності свого попередника, але й Microsoft інтегрувала ChatGPT у свою систему пошуку Bing, а Google випустив свій GPT-чат-бот Bard. Тепер GPT-4 може генерувати набагато більш точні, креативні відповіді та брати участь у все більш широкому спектрі заходів, таких як здача адвокатського іспиту [88].

Створення ШІ та його використання в різних видах діяльності людини, в тому числі в Інтернеті, викликало багато дискусій в різних галузях, а саме соціологічну філософію, машинобудування, вимірювальну техніку, ІКТ робототехніку, медицину, освіту, військову справу і право.

Поява сучасного ШІ є результатом кількох ключових віх, які значно вплинули на його розвиток та застосування:

1. Теоретичні основи виникнення ШІ. Формування основних теоретичних концепцій ШІ, зокрема блоків символного програмування, машинного навчання та теорії прийняття рішень. Розвиток цих ідей відкрив шлях для подальшого розвитку ШІ.

2. Перші успіхи в розвитку ШІ. Значні досягнення, такі як програма Logic Theorist, створена Алленом Ньюеллом і Гербертом Саймоном, яка вперше

продемонструвала можливість комп'ютерів вирішувати складні математичні проблеми.

3. Теорія нейромереж. Ідеї щодо Штучних нейромереж, які стали основою для подальшого розвитку і використання нейронної мережі в сучасному ШІ.

4. Експоненційний розвиток обчислювальної техніки. Прогрес в обчислювальній техніці, що дозволило використовувати більш потужні обчислювальні ресурси для вирішення складних задач ШІ.

5. Зростання обсягів легкодоступних даних. З появою Інтернету і цифрової трансформації зростає обсягів даних, що стали доступними для аналізу та обробки. Це сприяло розвитку способів машинного навчання, що стали підґрунтям для значної кількості сучасних досягнень ШІ.

6. Розробка алгоритмів глибокого навчання. Востаннє десятиліття було знаменним для розвитку глибокого навчання, що дозволило досягти нових висот у сферах комп'ютерного зору, опрацювання живої мови, ігор та інших областей.

Ці віхи в історії сприяли значному прогресу з розробок та застосуванню сучасного ШІ [105].

Існування численних наукових установ і колективів, незалежних вчених, що в основному працюють над конкретними проблемами ШІ, крім того міждисциплінарний характер наведених досліджень є причинами численних варіацій, що з'явилися у визначенні поняття «Штучний інтелект» [115, с. 2].

Сьогодні, а особливо близьке майбутнє, визначається тим, що впровадження в діяльність технологій ШІ більше поширюється та стає інтернаціональним, незважаючи на національні кордони. Спільне розуміння ШІ є ключовим для його оптимального застосування, керування та управління ним, розробки гармонізованих міжнародних і національних законів, а також реалізації скоординованої національної політики, спрямованої на встановлення сталого розвитку.

Історично склалося так, що американський інформатик Джон Маккарті вперше запропонував таке визначення ШІ на Дартмутській конференції в 1956 р., а саме: «Штучний інтелект - це розробка гнучких агентів, здатних адаптуватися до різноманітних ситуацій, раніше не відомих або не вивчених на досвіді, і досягати цілей, неможливих за допомогою загальноприйнятих систем комп'ютера» [86].

Відомий юридичний словник Блека визначає ШІ як «ПЗ, яке використовується для того, щоб змусити комп'ютери та роботів працювати ефективніше людей» [107].

Експертна група високого рівня зі Штучного інтелекту Європейська комісія (High-Level Expert Group on Artificial Intelligence European Commission) запропонувала таке визначення: «Штучний інтелект (ШІ) стосується систем, які демонструють інтелектуальну поведінку шляхом аналізу свого оточення та виконання дій – з конкретним ступенем незалежностей – для виконання поставлених цілей. Системи ШІ можуть бути повністю програмними, працювати у віртуальному світі (наприклад, помічники по голосу, програми для аналізування фото, пошукові системи, системи розпізнавання мовлення та рис обличчя), або бути вбудованими у пристрої (наприклад, роботи, автономні автомобілі, дрони чи програми Інтернету речей)» [98 с. 26].

У звіті Об'єднаного дослідницького центру Єврокомісії (Joint Research Centre - European Commission), як базовій роботі, що аналізує численні офіційні та наукові визначення терміну «Штучний інтелект», було сформульовано майже аналогічну версію визначення [98, с. 24]. Водночас у звіті зазначається, що «системи ШІ можуть застосовувати символічні правила або цифрові моделі і можуть адаптувати власну поведінку, проаналізувавши, як їхня попередня поведінка впливає на навколишнє середовище» [98, с. 16].

В ґрунтовному дослідженні групи експертів ШІ Організації економічного співробітництва та розвитку (ОЕСР) «Штучний інтелект у суспільстві» 2019 р. визначено систему Штучного інтелекту «як машинну систему, що може

здійснювати прогнозування, надавати рекомендації та приймати рішення, які впливатимуть на реальність чи віртуальний світ для досягнення певних цілей, заданих людиною» [88, с. 9].

Помічено, що людина з її здатністю розрізняти і судити, вищим розумінням і творчими здібностями перевершує машини з погляду використання інстинкту, здорового глузду і досвіду з життя при вирішенні завдань, формуванні рішень і вирішенні проблем [121].

Крім того, необхідно враховувати підходи до розуміння Штучного інтелекту, які є найбільш значущими і релевантними для цього дослідження:

- вчений у галузі інформатики Андреас Хольцінгер (Andreas Holzinger) зазначає, що «це найдавніша галузь комп'ютерно-технічних наук, що займається всіма аспектами створення систем, що імітують когнітивні функції, навчаються і мислять, як люди, задля вирішення реальних проблем» [71];

- професор інформатики Білл Томлінсон (Bill Tomlinson) визначає поняття, як «інтелект, що дає машинам і комп'ютерам імітувати або копіювати людські здібності» [76];

- литовський адвокат Паулюс Церка (Paulus Tserka) у своїх роботах описує поняття, як «Штучно створений інтелект, тобто програмні системи, які імітують людське мислення на ПК та інших пристроях, такі як система керування будинком, вбудовані в побутову техніку, роботи, автоматизований транспорт, літаки тощо» [40].

Існує низка визначень терміну «Штучний інтелект», які не узгоджуються між собою, оскільки охоплюють різні суб'єкти, об'єкти та цілі досліджень. Тим не менш, всі вони досить близькі в деяких важливих аспектах, тому існує потенційна можливість досягти взаєморозуміння даного питання [113].

Загалом, всі підходи до визнання і розуміння поняття «Штучний інтелект» можна розподілити на 2 групи:

- 1) Перша значень приписує «машинам», «механічним системам», «комп'ютерам», «програмним системам», «технологіям», «агентам» і т.д. певні

несуттєві характеристики, такі як здатність до адаптації; здатність досягати цілей; здатність здійснювати прогнози, створювати рекомендації та рішення, що мають вплив на будь-яке середовище; здатність сприймати це середовище; здатність групувати дані; здатність приймати ефективні рішення тощо [36].

2) Друга включає здатність виконувати когнітивні функції, вчитися і мислити як людина, реалізовувати сприйняття, розуміння, планування, дії, вчитися на рівні інтелекту, подібному до людського, виконувати корисні когнітивні завдання навіть краще за людину, імітувати або копіювати людські інтелектуальні здібності [42].

За останнє десятиліття багато країн визнали, що їхній майбутній успіх залежить від опрацювання великих об'ємів даних і можливостей ШІ. Багато приватних і державних організацій, науково-дослідних інститутів та академічних установ йдуть типовим шляхом у розробці та створенні продуктів ШІ. Спочатку вони нарощують власні інтелектуальні ресурси для прискорення отримання кінцевих результатів досліджень і вирішення бізнес-завдань [42]. Зазвичай, результати, отримані цими окремими командами розробників ШІ, не відповідають глобальним цілям або мають звужене застосування. Тоді команда розробників вступає в так званий «льодовиковий період», за час якого проект або припиняється, або дуже повільно пристосовується до нових задач: незабаром стає зрозуміло, що невеликі успіхи у створенні ШІ-технологій не мають довгострокових наукових і технологічних перспектив [96]. Лише глобальний, інклюзивний і керований державою підхід відкриває більше можливостей для розвивання сучасних ІКТ задля будування перспектив [1]. Все це набуває першочергового значення в контексті необхідності вжиття різноманітних заходів для подолання деградації цивілізації.

Місія Штучного інтелекту полягає в створенні умов для значного підвищення ефективності соціальної та виробничої діяльності в суспільстві, забезпечуючи прийняття якісних (оптимальних) рішень без негативних впливів людського фактору. Щодо мети впровадження ШІ варто відзначити висновок

про те, що останніми роками формуються все більш складні умови для діяльності людини [43].

Одночасно дедалі частіше з'являються повідомлення про те, що впровадження технологій ШІ зможе посилювати ефективність ІТ-комплексів і систем [41]. Ці повідомлення стосуються автономно керованих транспортних засобів (літаків, кораблів, автомобілів), різноманітних промислових, медичних, будівельних, освітніх, професійних, побутових, військових та роботів спеціального призначення [114].

Більшість думок щодо цілей застосування ШІ є однотипними [див. Додаток №1, с.122].

В загальному метою впровадження ШІ є забезпечення прийняття якісних (оптимальних) рішень та їх подальша ефективна реалізація [див. Додаток № 2, с. 123].

Швидке впровадження ШІ у фінансовому та економічному секторах спонукає глобальні інституції, такі як Міжнародний валютний фонд і Світовий банк, розробляють особисті стратегії для реакції на сучасні трансформаційні цифрові процеси [38].

Що стосується цифрової трансформації, ШІ має вужче значення, ніж загальне ПЗ, яке імітує людську поведінку. У цьому контексті ШІ означає алгоритм із чітко визначеним набором вхідних даних, розроблений для попередньо визначеного завдання, який зазвичай можна побачити в програмах САПР, які працюють із певними параметрами [23].

Вплив ШІ на цифрову трансформацію багатогранний. За своєю суттю ШІ функціонує, обробляючи великий обсяг даних і генеруючи практичні висновки. Використання ШІ вже давно поширене в багатьох галузях, але останніми роками його здатність автоматично пов'язувати складні алгоритми з великими обсягами даних набрала обертів. Це поширилося і на міжнародні відносини.

З урахуванням норм реалій ШІ відіграє помітну роль у міжнародному процесі цифровізації з наступними перспективами:

1. Використовується для розробки різноманітних нових продуктів та послуг, що змінюють способи, спілкування, праці, вироблення і їх споживання.
2. Допомагає автоматизувати безліч рутинних завдань і оптимізувати процеси у бізнесі, громадському секторі та інших сфер діяльності.
3. Допомагає в аналізі значних об'ємів даних, які є головною складовою цифровізації. Також може знаходити закономірності, тренди та взаємозв'язок в даних, що допомагає у прийнятті кращих стратегічних рішень.
4. Використовується для виявлення та запобігання кіберзлочинності та інших загроз цифровізації. Він допомагає виявляти аномальні збої в мережевому трафіку, захищаючи тим самим важливі інфраструктури та дані.
5. У міжнародному контексті, розвиток та застосування ШІ вимагає співробітництва між країнами. Це включає обмін досвідом, розробку міжнародних стандартів та правил щодо етики, безпеки та регулювання використання ШІ [11].

Отже, можна відзначити, що Штучний інтелект (ШІ) активно впливає на економічні, соціальні та технічні аспекти життя, сприяючи автоматизації процесів, оптимізації ресурсів та підвищенню продуктивності. Використання ШІ допомагає аналізувати великий об'єм інформації, знаходити закономірності та створювати стратегії, що підвищують ефективність у бізнесі, медицині, освіті й інших сферах. Однак розвиток цих технологій несе ризики, зокрема етичні питання щодо прозорості механізмів, захищеності даних і впливу на ринок праці.

Історичний розвиток ШІ свідчить про постійні зміни в його концепції та застосуванні. Від перших програм доведення теорем і шахових систем до сучасних нейромереж і моделей генерації, як GPT, технологія пройшла шлях від теоретичних розробок до практичних застосувань. Проте залишається проблема неоднозначності визначення ШІ, що часто залежить від сфери його використання: від вузькоспеціалізованих алгоритмів до систем, які імітують когнітивні здібності людини.

Перспективи використання ІІІ великі, проте вони вимагають відповідального підходу до регулювання та етики. Міжнародна співпраця у створенні стандартів для ІІІ необхідна для гармонізації його використання, запобігання зловживанням і гарантування сталого розвитку. Водночас, важливо враховувати культурні, соціальні та економічні контексти для досягнення балансу між технічним прогресом і потребами суспільства.

1.2. Стан джерельної бази дослідження

Сфера досліджень ІІІ наразі є розмитою та невизначеною, але настільки популярною, що її обговорюють не тільки дослідники та практики, а й бізнесмени, соціологи, державні установи, обізнані журналісти та інші пов'язані сторони, що мають власні погляди та розуміння цієї технології. Участь експертів з різним рівнем спеціалізації з різних сфер у таких важливих для суспільства дебатах ще більше ускладнює формування чіткого уявлення про ідентичність ІІІ та його ключові характеристики [51].

Існує величезна кількість літератури про вплив програм ІІІ в різних областях. Однак, щоб перейти до перехресного дослідження впливу ІІІ, потрібно зробити крок назад і з'ясувати, з якою технологією сучасне суспільство має справу.

Однією з відмінних характеристик ІІІ є широта його застосувань. У науковій літературі технології, які піддаються широкому застосуванню, називаються «технологіями загального призначення». У такому розумінні ІІІ можна порівняти з паровою машиною, електрикою та двигуном внутрішнього згоряння. Термін, який прийнятий, щоб передати природу ІІІ, — це «системна технологія», причому слово «система» тут стосується багатьох різних технологій, які входять до складу ІІІ та пов'язані з ним, а також його впливовості на суспільство.

Характеристика ШІ як системної технології має безпосередні наслідки для того, як варто розглядати дане питання.

Його вплив зараз є темою великої кількості літератури, а також незліченних принципів і статутів. Нещодавно проведений перелік містить понад 300 наборів етичних кодексів і настанов, що стосуються Штучного інтелекту. Яскраві приклади включають ті, які були розроблені Групою експертів високого рівня Європейської комісії з питань ШІ (European Commission's High-Level Expert Group on AI - AI HLEG) [62], ЮНЕСКО [67] та AI Now Institute [108].

Багато публікацій пов'язують вплив технології з такими цінностями, як пояснюваність, прозорість, відсутність дискримінації, конфіденційність, автономія та відповідальність. У той же час небезпечно намагатися звести вплив ШІ до списку суспільних цінностей, оскільки це суперечить його динамічному входженню в наше суспільство.

Якщо ШІ є системною технологією, то її вплив на суспільні цінності не можна просто звести до списку ефектів. На це є кілька причин. По-перше, як системна технологія ШІ все частіше буде використовуватися в суспільстві. Крім того, оскільки ми все ще перебуваємо на ранніх стадіях його розробки, жоден список не може бути нічим іншим, ніж тимчасовим. По-друге, технологія має впливати не лише на «специфічні для Штучного інтелекту» цінності, згадані вище, але й на ті, що є центральними для контексту, у якому сама технологія застосовується. Якщо ШІ можна використовувати в певному контексті, він потенційно може вплинути на всі суспільні цінності, пов'язані з цим контекстом.

Історія ШІ описана в роботі Алана Тюрінга в 1950 році та Дартмутської конференції 1956 року, де Джон Маккарті ввів сам термін [108].

Важливими етапами стали створення програми Logic Theorist, експертної системи MYCIN у 1970-х і алгоритму backpropagation у 1986 році, що поклав початок сучасному глибинному навчанню [98].

Крім того, 1990-ті роки ознаменували прорив у обробці природних мов і комп'ютерному зорі, а 2000-ні привели до розвитку об'ємних даних і генеративного ШІ, таких як GPT і DALL-E. [102].

Сьогодні ШІ охоплює нейронні мережі, що революціонізували аналіз даних, розпізнавання мовлення й образів, та використовується у широкому спектрі галузей.

Аналіз, статей і публікацій щодо ШІ, дуже важливий, оскільки він проливає світло на те, що зараз відбувається, і інформує про дискусію, яка зараз ведеться. Однак небезпека заключається в тому, що якщо б такий аналіз тлумачити як всеосяжний, то це може призвести до помилкового уявлення про те, що впливом ШІ можна керувати, доки захищено відповідні цінності.

Нарешті, важливо визнати, що сама концепція «впливу» вводить в оману. Якщо ми розглядаємо суспільство та його основні цінності як статичні, ми схильні розглядати ШІ як зовнішнє явище, яке потенційно може підірвати ці цінності – і дебати щодо ШІ справді часто формуються в таких термінах [86].

Сьогодні в зарубіжній практиці видавничу діяльність здійснює 19 спеціалізованих журналів, які присвячені Штучному інтелекту (див. Додаток 4, ст. 125) [64].

Ці спеціалізовані журнали були обрані з рейтингу Scimago Journal Rankings для ШІ (SJR 2020) і рекомендованого списку журналів Китайської комп'ютерної федерації (CCF 2019). Було обрано лише журнали найвищого рівня, які зосереджуються на основних технологіях Штучного інтелекту. Також відбулося вилучення журналів, які наголошували на функціональних застосуваннях ШІ (наприклад, журнал Artificial Intelligence in Medicine не був обраний).

З 19 вибраних журналів вищого рівня всі є міжнародними, 11 визначені як Scimago, так і Китайською комп'ютерною федерацією, тоді як інші вісім належать Scimago. Всі вони знаходяться в WoS у предметній категорії

«Штучний інтелект» [64]. Саме ці журнали та висвітлюють основні питання діяльності ШІ в умовах сьогодення, розкривають нові можливості ШІ та проведення дискусій про його подальше впровадження.

Більш детального розголосу Штучний інтелект отримує в нормативному визнанні.

У ЄС 13.03.2024 р Європейський парламент проголосував за закон про Штучний інтелект. Очікується, що закон набуде чинності в найближчі місяці, а після 2-річного пільгового періоду почне застосовуватися приблизно в червні. 2026 р. [26]. Також, Європейський Союз використовує Регламент Європейського Парламенту та Ради щодо гармонізованих правил у сфері Штучного інтелекту (Artificial Intelligence Act), а також вносить зміни до деяких інших актів Союзу [34].

Окрім Закону про Штучний інтелект, інші важливі закони у сфері Штучного інтелекту знаходяться в стадії розробки. Зокрема, запропонована Директива про відповідальність за Штучний інтелект розроблена Європейською комісією для адаптації законодавства до особливостей цифрової економіки. спрямована на роз'яснення цивільної відповідальності щодо збитків, спричинених системами ШІ [73].

Рада Європи для регулювання Штучного інтелекту застосовує рамкову конвенцію, міжнародний договір, відомий як Рамкова конвенція про Штучний інтелект, права людини, демократію та верховенство права [74].

Хоча в США немає єдиного регулювання ШІ, вони встановили численні рекомендації та рамки для управління сектором на федеральному рівні. Існуючий правовий ландшафт охоплює розпорядження, як-от: «Збереження американського лідерства в ШІ» 2019 р. [73], «Закон про сприяння використанню надійного Штучного інтелекту у федеральному уряді» 2020 р. [65], «законопроект 311 Асамблеї Каліфорнії» 2021 р. [95].

В Україні основним джерелом для дослідження Штучного інтелекту стала Концепція розвитку Штучного інтелекту в Україні, затверджена Розпорядженням КМУ від 02.12.2020 року № 1556-р. [15].

За результатами дослідження можна виокремити окрему групу дослідників, які працювали над безпосереднім визначенням і поняття ШІ.

Розвиток ШІ супроводжувався спробами теоретично окреслити його сутність. Однією з першопочаткових спроб стала робота Джона Маккарті, який на Дартмутській конференції в 1956 році ввів поняття ШІ [86]. У його визначенні ШІ був представлений як розробка "гнучких агентів", здатних адаптуватися до раніше невідомих обставин і досягати цілей. Саме ця конференція заклала основу для теоретичних і практичних досліджень у цій сфері.

Іншим важливим внеском стало тестування можливостей машинного інтелекту через так званий тест Тюрінга, запропонований британським математиком Аланом Тюрінгом. Його метою було визначити, чи здатна машина мислити, як людина. Хоча цей підхід викликав численні дискусії, він залишився актуальним у сучасних дослідженнях.

У ширшому розумінні ШІ визначають як системи, здатні демонструвати інтелектуальну поведінку через аналіз даних і автономне прийняття рішень. Таке розуміння було запропоноване Групою експертів високого рівня Європейської комісії у звіті 2020 року. Там зазначалося, що ШІ може функціонувати як у віртуальному середовищі (голосові помічники, алгоритми розпізнавання облич тощо), так і у фізичних пристроях (автономні автомобілі, дрони).

Андреас Хольцінгер [71] у своїй роботі наголошує, що ШІ імітує когнітивну функцію людини, включаючи навчання, планування й ухвалення рішень. Литовський дослідник Паулюс Церка розглядає ШІ як програмні системи, що імітують людське мислення, застосовуючи їх до автоматизованого транспорту, побутової техніки й роботів.

Існують дві ключові групи визначень ШІ. Перша зосереджена на технічних аспектах, описуючи системи як здатні адаптуватися й аналізувати середовище для прийняття рішень. Друга група фокусується на когнітивному підході, зазначаючи, що ШІ імітує розумові процеси, властиві людині. Такий поділ відображає багатовимірність поняття ШІ й демонструє складність його застосувань.

Серед найвизначніших праць у даній сфері необхідно вкредити звіти Єврокомісії, дослідження OpenAI та статті Андреаса Хольцінгера. Всі вони наголошують, що визначення ШІ залежить від контексту, в якому використовується, та від специфіки задач, які він виконує [88].

Окрему віху в дослідженні складають загрози пов'язані з використанням ШІ.

ШІ відкриває безліч можливостей, проте водночас породжує значні загрози, які дослідники активно обговорюють у своїх працях. Однією з найгостріших проблем є питання безпечності інформації та конфіденційності. Сем Альтман, генеральний директор OpenAI, висвітлює ризики, які пов'язані з несанкціонованим вилученням даних у мовних моделях, таких як ChatGPT. Він наголошує, що попри заходи безпеки, системи ШІ можуть бути зламані, що створює серйозні ризики для конфіденційності користувачів.

Ще одна загроза пов'язана з можливістю використання штучного інтелекту для створення шкідливих програм. За словами дослідників Колумбійської бізнес-школи, генеративні моделі, такі як ChatGPT, можуть використовуватися для автоматизації створення вірусів і ботів. Це значно знижує поріг входження для кіберзлочинців, які раніше потребували технічної експертизи.

Проблема дезінформації також стала центральною у сучасних дебатах про ШІ. Організація NewsGuard зазначає, що з 2022 року кількість сайтів, які публікують фейкові новини, створені за допомогою генеративного ШІ, зростає.

на 1000%. Це викликає занепокоєння, адже такі системи здатні автоматизувати створення переконливих, але неправдивих статей.

Особливу увагу дослідники приділяють проблемі глибинних фейків (deepfakes). Наприклад, у 2021 році хакери використали технологію deepfake для шахрайства, видаючи себе за керівника компанії, що призвело до фінансових втрат. Подібні випадки демонструють, як ШІ може використовуватися для маніпуляцій і підриву довіри [114].

Ризики ШІ поширюються й на фізичну безпеку. Дослідження показують, що автоматичний транспорт або виробниче обладнання, під'єднане до ШІ, може бути вразливим до атак. Маніпуляції з інформацією можуть призвести до збоїв у роботі систем і створити загрозу для життя людей.

Важливим напрямком досліджень є вплив ШІ на ринок праці. Багато науковців застерігають, що автоматизація робочих процесів може призвести до скорочення робочих місць у таких секторах, як виробництво, логістика та навіть програмування. Водночас зростає попит на спеціалістів, які здатні розробляти й обслуговувати системи ШІ, що змінює структуру зайнятості.

Таким чином, основні загрози ШІ стосуються безпеки, етики й соціальних аспектів. Хоча дослідження, такі як роботи Альтмана чи аналітика NewsGuard, демонструють реальні ризики, вони також наголошують на необхідності розробки системного підходу до регулювання технологій.

Проблематика правового регулювання Штучного інтелекту є однією із центральних питань сучасної наукової та юридичної дискусії. Важливість цього аспекту підкреслюється численними дослідженнями, присвяченими нормативно-правовим викликам, які виникають через використання ШІ. Відомий внесок у цю сферу зробила Європейська комісія, яка ще у 2021 році запропонувала Регламент про Штучний інтелект (Artificial Intelligence Act), направлений на встановлення гармонійних зв'язків та правил використання ШІ в країнах ЄС [34]. Регламент розділяє всі системи ШІ на категорії ризику

(недопустимий, високий, обмежений і мінімальний), що дає змогу оцінювати потенційну небезпеку та відповідно регулювати використання таких систем.

Європейські ініціативи стали взірцем для великої кількості країн. До прикладу, у США, попри відсутність єдиного федерального законодавства, існують численні рамкові документи, такі як «Закон про сприяння використанню надійного ШІ» (2020 рік) [65]. Він пропонує принципи етичного використання ШІ в державному управлінні, зокрема прозорість і відповідальність за наслідки.

В Україні важливим кроком стало прийняття Концепції розвитку Штучного інтелекту [13], затвердженої ЗМУ у 2020 році. Документ формує головні напрями держполітики в ШІ, включаючи створення інфраструктури для досліджень, впровадження технологій у держоргани та зростання обізнаності населення про ШІ.

Серед найвагоміших праць у цій галузі можна виділити роботи міжнародних організацій, таких як Рада Європи, яка застосовує рамкову Конвенцію з питань ШІ та прав людини. Документ наголошує на необхідності захисту основних прав і свобод у процесі розробки й впровадження ШІ.

Велика увага надається питанням відповідальності за наслідки використання ШІ. Наприклад, у запропонованій Директиві ЄС [45] щодо відповідальності за ШІ йдеться про адаптацію існуючого законодавства до особливостей цифрових технологій. Особливий акцент зроблено на цивільній відповідальності за шкоду, завдану автономними системами.

Хоча існує безліч підходів до регулювання ШІ, дослідники сходяться на думці, що глобальна співпраця є ключовою для ефективного управління цими технологіями. Внесок Європейської комісії, Ради Європи, США та України формує основу для подальшого розвитку нормативно-правової бази.

Оскільки ШІ є технологією глобального масштабу, його регулювання вимагає співпраці між країнами. У цьому контексті міжнародні ініціативи

відіграють важливу роль у встановленні стандартів використання та розробки ШІ.

Серед більш впливових ініціатив можна виділити Глобальну ініціативу з управління ШІ (Global Partnership on AI – GPAI), засновану в 2020 році за підтримки Організації економічного співробітництва та розвитку (ОЕСР) [88]. Ця організація об'єднує експертів з понад 20 країн для розробки рекомендацій, які забезпечують етичне використання ШІ. GPAI зосереджує свою увагу на таких напрямках, як прозорість алгоритмів, запобігання упередженням у даних і захист конфіденційності.

Важливою є також діяльність ЮНЕСКО, яка у 2021 році прийняла рекомендації щодо етичних аспектів ШІ [78]. Документ включає принципи недискримінації, екологічної сталості та поваги прав людей. Це перший глобальний інструмент, що регулює застосування ШІ в соціально-етичному контексті.

Додатково, Рада Європи працює над розробкою Конвенції з питань етики й відповідальності у використанні ШІ. У її тексті наголошується, що ШІ повинен сприяти розвитку демократії, а не загрожувати їй.

У рамках об'єднань, таких як Європейський Союз, відбувається активна робота над створенням єдиного підходу до регулювання. Прийняття Artificial Intelligence Act [52,101] стане важливим кроком у цьому напрямку. У США жодна ініціатива не має настільки широкого масштабу, але активно діють рекомендації Національного інституту стандартів і технологій (NIST) щодо управління ризиками в системах ШІ.

На глобальному рівні відчутний вплив мають також приватні організації. Наприклад, OpenAI, заснована Ілоном Маском і Семом Альтманом [84], акцентує увагу на необхідності розробки безпечного ШІ. Їхній фокус на довгостроковій перспективі розвитку ШІ передбачає баланс між інноваціями та етикою.

Міжнародні ініціативи демонструють, що вирішення викликів ШП можливе лише через скоординовану співпрацю між державами, організаціями та приватним сектором. Однак багато дослідників зазначають, що регулювання залишається нерівномірним, що потребує посилення глобальних зусиль.

РОЗДІЛ 2 ВИКЛИКИ ТА НАСЛІДКИ ВИКОРИСТАННЯ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

2.1. Штучний інтелект та кібербезпека: ризики та можливості

Нові технологічні розробки в галузі Штучного інтелекту захопили світ штурмом, спонукаючи уряди домагатися стратегічної переваги, а технологічні компанії – розробляти та комерціалізувати нові системи ШІ.

Нові програми ШІ потенційно можуть принести численні переваги суспільству, але також створити серйозні наслідки для безпеки, починаючи від національної безпеки, дестабілізації демократії та закінчуючи масштабним економічним зривом. Наслідки ШІ для кібербезпеки численні та постійно розвиваються, причому суб'єкти загрози використовують ці нові технології на свою користь, розширюючи свої можливості для кібератак.

ШІ призведе до еволюції та вдосконалення існуючих тактик, методів і процедур, а також знизить бар'єр доступу для кіберзлочинців, скоротивши технічні знання, необхідні для здійснення кібератак. Соціальна інженерія також посилюється завдяки новим великим мовним моделям (LLM), за допомогою яких суб'єкти загроз створюють дедалі складніші фішингові кампанії [28].

У міру розвитку технології різницю між синтетичними носіями та контентом, створеним людьми, стає все важче помітити навіть для технологій виявлення, що робить дідфейки (deepfake) більш реалістичними, цілеспрямованими та небезпечними, ніж будь-коли раніше.

Нещодавно гонконгська філія транснаціональної компанії постраждала від шахрайства з підтримкою deepfake. Зловмисники застосували технологію deepfake, щоб маскуватися під фінансового директора компанії під час відеоконференції, спонукавши одного співробітника сплатити кошти.

Зростаючі загрози використання ШІ, не повинні затьмарювати великі переваги, які нові моделі можуть принести кібербезпеці. Оскільки ці

інструменти використовуються зловмисниками в кіберпросторі, кіберзахисники можуть, у свою чергу, використовувати ШІ для покращення можливостей кібербезпеки [28].

Сьогодні очікується, що ризики Штучного інтелекту для кібербезпеки швидко зростуть, оскільки інструменти ШІ стануть дешевшими та доступнішими. Варто виокремити наступні ризики ШІ в кібербезпеці:

1. Оптимізацію кібератак. Експерти вважають, що зловмисники можуть використовувати генеративний Штучний інтелект і великі мовні моделі для проведення атак з небаченою швидкістю і складністю. Вони застосовують ці технології для знаходження нових способів підризу безпеки в хмарах і використання геополітичної напруженості для проведення передових атак. Крім того, зловмисники можуть удосконалювати свої методи програм-вимагачів і фішингових атак за допомогою генеративного Штучного інтелекту.

2. Автоматизоване шкідливе ПЗ. Такий Штучний інтелект, як ChatGPT, чудово справляється з точним обчисленням чисел. За словами професора Колумбійської бізнес-школи Огеда Нетцера, ChatGPT уже може «досить добре писати код». Експерти відзначають, що в найближчому майбутньому це може допомогти розробникам програмного забезпечення, програмістам або зникне потреба в цих професіях [30].

Хоча таке ПЗ, як ChatGPT, має певний захист, щоб запобігти створенню шкідливого коду користувачами, експерти можуть використовувати розумні методи, щоб обійти його та створити зловмисне програмне забезпечення. Це може бути вершиною айсберга. Майбутні інструменти на базі Штучного інтелекту можуть дозволити розробникам із початковими навичками програмування створювати автоматизоване зловмисне програмне забезпечення, як просунутий шкідливий бот. Зловмисний бот може викрадати дані, заражати мережі та атакувати системи практично без втручання людини.

3. Фізична безпека. Оскільки все більше систем, таких як автономний транспорт, виробниче та будівельне обладнання та медичні системи,

використовують ШІ, ризики для фізичної безпеки можуть зростати. Наприклад, справжній безпілотний автомобіль на основі Штучного інтелекту, який зазнав порушення кібербезпеки, може призвести до ризику для фізичної безпеки його пасажирів. Так само зловмисник може маніпулювати набором даних для інструментів технічного обслуговування на будівельному майданчику, щоб створити небезпечні умови.

4. Ризики конфіденційності ШІ. Помилка, яка була незручною для генерального директора OpenAI Сема Альтмана, викликала витік фрагментів історії чату інших користувачів у ChatGPT. Хоча помилку було виправлено, існують інші можливі ризики конфіденційності через величезну кількість даних, які обробляє Штучний інтелект. Наприклад, хакер, який зламує систему ШІ, може отримати доступ до різних видів конфіденційної інформації.

Система ШІ, розроблена для маркетингу, реклами, профілювання або стеження, також може загрожувати конфіденційності даних та приватності користувачів. У деяких країнах технологія профілювання ШІ вже допомагає державам втручатися в конфіденційність персональних даних [30].

5. Крадіжка моделей ШІ. Існують певні ризики викрадення моделі Штучного інтелекту через мережеві атаки, методи соціальної інженерії та використання вразливостей суб'єктами загроз, такими як спонсоровані державою агенти, інсайдерські загрози, як-от корпоративні шпигуни, та звичайні комп'ютерні хакери. Викрадені моделі можна маніпулювати та модифікувати, щоб зловмисники здійснювали різні зловмисні дії, посилюючи ризики ШІ для суспільства.

6. Маніпулювання даними та отруєння даними. Хоча ШІ є потужним інструментом, він може бути вразливим до маніпулювання даними. Зрештою, ШІ залежить від своїх навчальних даних. Якщо дані змінено або отруєно, інструмент на основі ШІ може призвести до неочікуваних або навіть шкідливих результатів.

Теоретично зловмисник може отруїти навчальний набір даних шкідливими даними, щоб змінити результати моделі. Зловмисник також може ініціювати більш витончену форму маніпуляції, яка називається ін'єкцією упередженості. Такі атаки можуть бути особливо шкідливими в таких галузях, як охорона здоров'я, автомобільна промисловість і транспорт [30].

7. Уособлення. Щоб побачити, як інструменти на основі Штучного інтелекту допомагають кінематографу, можемо навести приклад документального фільму «Roadrunner», в якому голос покійного знаменитого шеф-кухаря Ентоні Бурдена створювався за допомогою аудіо, згенерованого Штучним інтелектом, і легко обманював глядачів. Подібним чином актор-ветеран Гаррісон Форд переконливо постарів на кілька десятиліть завдяки силі Штучного інтелекту в «Індіані Джонсі та циферблаті долі».

8. Більш витончені атаки. Як уже зазначалося, загрози можуть використовувати Штучний інтелект для створення вдосконаленого зловмисного програмного забезпечення, видавати себе за інших для шахрайства та модифікувати навчальні дані ШІ. Вони можуть автоматизувати фішингові атаки, атаки зловмисного програмного забезпечення та крадіжку облікових даних за допомогою ШІ. Крім того, Штучний інтелект може допомогти обійти системи безпеки, такі як програмне забезпечення для розпізнавання голосу, під час атак, відомих як змагальні атаки.

9. Шкода репутації. Організація, яка використовує ШІ, може постраждати від репутації, якщо технологія не працює або зазнає порушення кібербезпеки, що призводить до втрати даних. Такі організації можуть зіткнутися зі штрафами, цивільними покараннями та погіршенням відносин із клієнтами [30].

Всі неведені ризики використання ШІ спрямовані на реалізації чотирьох напрямків злочинних дій: дезінформація, кібератаки, глибинні фейки та атаки на критичну інфраструктуру.

В умовах сьогодення дезінформація є одним з важливих викликів, підкреслюючи, що швидкий розвиток технологій також створює нові проблеми або погіршує існуючі. Штучний інтелект автоматизує процес створення фейкових новин, що призводить до збільшення обсягу веб-контенту, який імітує справжні статті, але насправді поширює дезінформацію про вибори, війни та стихійні лиха. За даними NewsGuard, організації, яка моніторить дезінформацію, з травня 2022 кількість веб-сайтів, що публікують неправдиві статті, створені Штучним інтелектом, зросла більше ніж на 1000 відсотків — з 49 до понад 600 сайтів [122].

Раніше для пропагандистських операцій використовувалися армії низькооплачуваних працівників або добре організовані розвідувальні служби для створення сайтів, які здавалися б законними. Але Штучний інтелект спрощує цей процес, дозволяючи майже будь-кому — від агентів розвідки до звичайних підлітків — створювати подібний контент, який іноді важко відрізнити від справжніх новин [102].

Глибинні фейки, які отримали назву Deepfake — це форма Штучного інтелекту (AI), яка може використовуватися для створення переконливих фальшивих зображень, звуків і відео. Термін «deepfake» поєднує концепцію глибокого навчання з чимось фейком. Deepfake копіює фальшиві зображення та звуки та зшиває їх разом за допомогою алгоритмів машинного навчання. У результаті створює людей і події, яких не існує або не відбулося насправді.

Технологія Deepfake переважно використовується для деструктивних цілей, наприклад для введення громадськості в оману шляхом поширення неправдивої інформації чи пропаганди. Наприклад, підроблені відео можуть показувати світового лідера чи знаменитість, які говорять те, чого насправді не було.

Технологію Deepfake можна використовувати для багатьох цілей, зокрема:

- Шахрайство та обман. Кіберзлочинці можуть використовувати технологію deepfake для створення шахрайства, неправдивих заяв і містифікацій, які підривають і дестабілізують організації.

- Порнографія знаменитостей. Головною загрозою, яку створює deepfake, є порнографія без згоди, на яку припадає до 96% deepfake в Інтернеті. Технологія Deepfake також використовується для створення містифікацій порно з помстою.

- Маніпулювання виборами. Відео Deepfake використовувалися для розповсюдження фейкових відео світових лідерів, таких як Дональд Трамп і Барак Обама, Володимир Зеленський, а це викликає стурбованість, що вони можуть бути використані для маніпуляцій на виборах [48].

- Соціальна інженерія. Технологія Deepfake використовувалася в шахрайстві соціальної інженерії, коли аудіо-deepfake змушували людей повірити, що вони розмовляють з довіреними особами. Наприклад, генерального директора британської енергетичної компанії обманом змусили повірити, що він розмовляє з виконавчим директором материнської компанії в Німеччині. Голос deepfake видавав себе за виконавчого директора та переконав генерального директора перерахувати 220 000 євро на банківський рахунок передбачуваного угорського постачальника.

- Автоматизовані атаки дезінформації. Deepfake також можна використовувати для поширення автоматизованих атак дезінформації, таких як теорії змови та неправильні теорії про політичні та соціальні проблеми. Досить очевидним прикладом такого використання дипфейку є фальшиве відео засновника Facebook Марка Цукерберга, який стверджує, що має «повний контроль над даними мільярдів людей» завдяки Spectre, вигаданій організації в романах і фільмах про Джеймса Бонда.

- Крадіжка особистих даних і фінансове шахрайство.

Діпфейки можна помітити, розпізнавши незвичайну активність або неприродні рухи, зокрема:

✓ Неприродні рухи очей. Відсутність руху очей є хорошою ознакою глибоких фейків. Відтворити природний рух очей складно, оскільки очі людей зазвичай стежать за людиною, з якою вони розмовляють, і реагують на неї.

✓ Відсутність кліпання. Відтворити природну людську дію регулярного кліпання за допомогою технології deepfake важко.

✓ Неприродні вирази обличчя та морфінг обличчя - коли обличчя просто зливаються з одного зображення на інше. Це зазвичай призводить до незвичних або неприродних виразів обличчя [114].

✓ Неприродна форма тіла. Якщо тіло людини не має природної форми, то, швидше за все, це підробка. Технологія Deepfake здебільшого фокусується на обличчях, а не на всьому тілі, що призводить до неприродних форм тіла.

✓ Неприродне волосся. Фальшиві зображення не можуть створити реалістичні індивідуальні характеристики, наприклад кучеряве чи зіпсоване волосся.

✓ Аномальні кольори шкіри. Deepfakes не в змозі відтворити природні кольори зображень і відео. Це призводить до того, що вони демонструють аномальний колір шкіри.

✓ Незручне положення голови та тіла. Зображення Deepfake часто демонструють непослідовне або незграбне положення голови та тіла. Прикладами цього є різкі рухи та спотворені зображення, коли люди рухаються або повертають голову [114].

✓ Непослідовне положення обличчя. Зображення Deepfake часто демонструють непослідовне або незграбне положення голови та тіла. Прикладами цього є різкі рухи та спотворені зображення, коли люди рухаються або повертають голову.

✓ Дивне освітлення або зміна кольору. Подібно до причин неприродних тонів шкіри, глибокі підробки також схильні до знебарвлення, неправильних тіней і незвичайного освітлення.

✓ Погана синхронізація губ. У відео Deepfake, ймовірно, буде синхронізація губ, яка не співпадає зі словами, які вимовляють люди на відео [114].

Також, хакери використовують методи, керовані Штучним інтелектом, щоб зламати найнадійніші програми кіберзахисту. Ці кібератаки за допомогою Штучного інтелекту змінюють ландшафт кібербезпеки, тому вкрай важливо прийняти комплексну програму кіберзахисту, яка включає цілісний захист. Кібератаки, керовані Штучним інтелектом, передбачають використання вдосконалених алгоритмів машинного навчання для виявлення вразливостей, прогнозування моделей і використання слабких місць. Ефективність і швидкий аналіз даних дозволяють хакерам отримати тактичну перевагу, що призводить до швидкого вторгнення або знищення. Традиційних методів кібербезпеки вже недостатньо для боротьби зі складними атаками, оскільки кібератаки ШІ адаптуються та розвиваються в режимі реального часу.

ШІ відіграє ключову роль у захисті систем критичної інфраструктури. AI може посилити безпеку критичної інфраструктури, використовуючи передові механізми виявлення загроз і реагування. Системи кібербезпеки на основі Штучного інтелекту постійно перевіряють мережеву активність, системні записи та поведінку користувачів на предмет можливих загроз і відхилень. Ці системи використовують алгоритми машинного навчання для аналізу величезних обсягів даних, виявлення шаблонів, пов'язаних із добре відомими та нещодавно виявленими кібератаками, і вжиття оперативних профілактичних заходів для зменшення ризиків [49].

Попри виявлені ризики не потрібно забувати й про нові можливості. AI стає потужним інструментом у боротьбі з кіберзагрозами, оскільки він може допомогти швидше виявляти, аналізувати та реагувати на зловмисні атаки [49].

Деякі переваги ШІ в кібербезпеці включають:

1. Запобігання кібератакам. Системи кібербезпеки Штучного інтелекту дозволяють проактивно керувати кіберризиками. Вони часто можуть

використовувати методи розпізнавання шаблонів, щоб розпізнавати шкідливі програми, програми-вимагачі та інші форми кібератак, а потім запобігати їм, перш ніж вони спричинять проблему. Вони навіть можуть передбачати майбутні кібератаки, розпізнаючи, коли зловмисний код було змінено хакерами або кіберзлочинцями, намагаючись уникнути заходів безпеки мережі. Технології, які наразі мають такі можливості, включають Darktrace RESPOND, CrowdStrike Falcon, Tenable та IBM Watson AI [31].

Системи Штучного інтелекту також можуть аналізувати трафік вхідної електронної пошти, щоб виявляти та перехоплювати фішингові атаки або сповіщати команду кібербезпеки, якщо одна з них успішна [12].

2. Покращення реагування на інциденти. Коли кібератаки стаються, системи Штучного інтелекту можуть допомогти вам реагувати швидше — навіть у режимі реального часу — для локалізації або усунення збитків шляхом швидкого аналізу даних про інциденти та надання результатів керівникам кібербезпеки. Він також продемонстрував потенціал для розвитку можливостей самовідновлення, які дозволять йому автоматично реагувати на загрози, вторгнення та злами.

Враховуючи, що одна кібератака коштує організаціям у середньому майже 5 мільйонів доларів США, кожна деталь гнучкості, яку система Штучного інтелекту додає до вашого реагування на інциденти, допоможе вам заощадити багато грошей [31].

3. Підвищення точності виявлення загроз. Системи Штучного інтелекту набагато менше схильні до людських помилок, ніж аналітики кібербезпеки, які використовують ручні методи. Посилення цих команд за допомогою Штучного інтелекту може допомогти їм зменшити рівень помилкових спрацьовувань або пропущених загроз і звільнити їхній час для більш стратегічної роботи та кращого та швидшого реагування на інциденти. Це також забезпечує вищий рівень надійності вашої програми кібербезпеки.

4. Аналіз великих обсягів даних. Системи Штучного інтелекту надзвичайно добре просіюють величезні масиви даних і позначають підозрілі моделі чи дії. З огляду на обсяги даних, які створюються сьогодні, командам майже неможливо переглянути їх усі вручну. Передача ваших журналів безпеки, брандмауера, виявлення вторгнень та інших даних ІТ-безпеки системі Штучного інтелекту може допомогти їй розпізнати звичайну поведінку у вашій мережі та виявити будь-які проблеми, як-от підозрілу активність, яка може свідчити про внутрішню загрозу або порушення даних.

Моделі Штучного інтелекту також можуть прискорити ваші програми кількісної оцінки кіберризиків, виконуючи кількісний аналіз ризиків у режимі реального часу, рекомендуючи заходи пом'якшення та допомагаючи вам відстежувати ключові показники ризику для проактивного керування кіберризиками [31].

5. Автоматизований безперервний контроль і тестування вразливостей. Системи Штучного інтелекту можуть використовуватися для автоматизації безперервного моніторингу та тестування елементів керування кібербезпекою, уразливостей і керування виправленнями у вашій організації. Виконання цієї роботи вручну - дуже трудомісткий процес. Дозволивши Штучному інтелекту виконувати це автоматично та безперервно, ви можете виявити та виправити будь-які прогалини в режимі реального часу та підтримувати готовність до аудиту в будь-який час.

6. Підвищення ефективності. ШІ може значно полегшити виконання таких трудомістких завдань, як розробка політики та процедурних документів. Такі системи, як ChatGPT, можуть дати кіберкомандам швидкий початок такого типу роботи, автоматично розробляючи схему або навіть першу чернетку цих документів [31].

7. Удосконалення культури кіберризиків і навчання з кібербезпеки. Системи Штучного інтелекту можна використовувати для створення більш реалістичних симуляцій кібератак, таких як спроби фішингу, на основі

реальних прикладів, які можна використовувати для підвищення обізнаності про кіберризики та закріплення належних практик кібергігієни у вашій організації.

8. Можливість масштабування. Системи Штучного інтелекту можна легко налаштувати та покращити для масштабування в міру того, як ваша організація та її програми кібербезпеки зростають, а також у міру того, як змінюється ландшафт ризиків кібербезпеки.

9. Інше використання. Вище наведено деякі з основних варіантів використання Штучного інтелекту в кібербезпеці, які ми зараз бачимо, але є ще багато, багато інших, зокрема: Прогнозування ризику порушення; Аутентифікація користувача; Виявлення шахрайства; Автоматизований аналіз загроз; Спам-фільтри [31].

В загальному використання Штучного інтелекту допоможе краще зрозуміти свої мережі та швидше виявляти потенційні загрози. Рішення на основі Штучного інтелекту можуть просіювати величезні масиви даних, щоб ідентифікувати аномальну поведінку та виявляти зловмисну активність, наприклад нову атаку нульового дня [29]. Штучний інтелект також може автоматизувати багато процесів безпеки, як-от керування виправленнями, що полегшує контроль за вашими потребами в кібербезпеці.

Системи кібербезпеки на основі ШІ забезпечують кращу точність і ефективність порівняно з традиційними рішеннями безпеки. Наприклад, Штучний інтелект може сканувати безліч пристроїв на наявність потенційних вразливостей за частку часу, який знадобився б операторам для виконання того ж завдання. Крім того, алгоритми Штучного інтелекту можуть розпізнавати шаблони, які важко помітити людському оку, що забезпечує більш точне виявлення зловмисної діяльності [29].

Як стверджував Джіту Патель із Cisco, «це чудовий час, щоб схилити терези на користь захисників» [28]. Ця думка широко поділяється серед технічної спільноти, як було продемонстровано на Мюнхенській конференції з

безпеки, де понад 15 провідних підприємств пообіцяли допомогти запобігти втручанню оманливого вмісту Штучного інтелекту в глобальні вибори цього року.

Крім того, індійсько-американський бізнесмен Сундар Пічаї з Google оголосив про новий напрямок роботи для посилення кіберзахисту шляхом прискорення виявлення та реагування груп. Корпорація Майкрософт очолює зусилля з виявлення та блокування використання зловмисниками служб Microsoft, а також запустила новий інструмент, який дозволяє користувачам використовувати цифровий підпис та автентифікувати носії. Meta оголосила про нові технологічні стандарти для ідентифікації та позначення контенту, створеного ШІ [28].

Технології Штучного інтелекту також можуть сприяти навчанню з кібербезпеки, чи то для навчання широкої громадськості, чи для підготовки наступного покоління кіберзахисників. Згідно з дослідженням ISC2 Cybersecurity Workforce Study, 2019 не вистачає біля 4 мільйонів фахівців з кібербезпеки [28].

Щоб допомогти усунути цю прогалину, у квітні 2023 року Форум запустив багатосторонню ініціативу під назвою «Подолання розриву в кібернавичках». Його мета полягає в тому, щоб створити стратегічну структуру кадрів у сфері кібербезпеки та розробити дії, які допоможуть людям увійти в робочу силу з кібербезпеки та процвітати в ній.

У змаганні за домінування Штучного інтелекту центральне місце займає складний зв'язок між кібербезпекою та цими передовими технологіями. Хоча Штучний інтелект підвищить ефективність кібератак, їхній вплив можна компенсувати перевагами використання технологій Штучного інтелекту для посилення механізмів кіберзахисту [28].

Отже, Штучний інтелект (ШІ) стрімко змінює технологічний ландшафт, створюючи можливості та виклики для безпеки. Його широке впровадження дозволяє оптимізувати процеси, але водночас посилює ризики кібератаки та

зловживання технологіями. Загрози включають використання deepfake для шахрайства, автоматизацію шкідливого ПЗ і маніпуляції даними. Сучасні атаки, керовані ШІ, стають складнішими, а розробка інструментів для злочинців спрацьовується через доступність генеративних моделей ризику, тому ШІ є потужним союзником у сфері кібербезпеки.

Системи на основі ШІ можуть активно виявляти загрози, автоматизувати аналіз даних і покращувати реагування на інциденти. Інструменти кіберзахисту допомагають запобігати атаці, підвищувати точність виявлення й ефективність моніторингу. Удосконалення технологій і навчання спеціалістів із кібербезпеки сприяють зниженню вразливості та створенню умов для безпечного розвитку цифрового суспільства.

2.2. Вплив Штучного інтелекту на права людини, демократію та верховенство права

Технології ШІ створюють нові можливості для забезпечення прав і свобод, зокрема шляхом автоматизації судових процесів, підвищення доступності до правосуддя та моніторингу порушених прав людини. Однак їх використання також несе серйозні ризики. Наприклад, алгоритми можуть посилювати дискримінацію через упередження в даних, а також ставити під загрозу приватності та вільного слова неконтрольоване збирання та аналіз персональної інформації.

Демократія й верховенство права також стикаються з новими викликами через поширення дезінформації, маніпуляцію громадською думкою та втручання у виборчі процеси за допомогою ШІ. Автономні системи прийняття рішень можуть зменшити прозорість і підзвітність державних і приватних організацій. Таким чином, інтеграція ШІ потребує збалансованого підходу, який забезпечує як використання його потенціалу, так і мінімізації загроз для прав людини і демократичних цінностей.

Багато систем або використання ШІ можуть одночасно впливати на права людини, демократію та правову державу, або негативно впливати на права однієї людини, тоді як позитивно впливають на права іншої.

Згідно з Європейської конвенції про захист прав людини піддається впливу ШІ: повага до людських цінностей - свобода особи, рівність, недискримінація та солідарність, соціальні та економічні права. Крім того, ШІ має значний вплив на демократію та правову державу [104].

Повага до людської гідності закріплена в Європейській конвенції з прав людини (ЄКПЛ) 1950 року через такі права, як право на свободу та безпеку (ст. 5), право на справедливий суд (ст. 6), право на захист від покарання без закону (ст. 7) і право на приватність, фізичну та психічну цілісність (ст. 8) [9]. Штучний інтелект може впливати на всі ці права, може підтримувати або посилювати існуючі упередження, що є критичним при його використанні в правоохоронних органах та судовій системі.

У ситуаціях, коли ставиться під загрозу фізична свобода або особиста безпека, наприклад, при передбачуваному поліцейському контролі, визначенні ризику повторного вчинення злочину та вироку, право на свободу та безпеку, поєднане з правом на справедливий суд, є вразливим [104].

Багато систем Штучного інтелекту та їх використання мають широкий і глибокий вплив на приватність. Проте слід зазначити, що хоча конфіденційність даних дійсно є важливим елементом, вплив ШІ на нашу конфіденційність виходить далеко за рамки наших персональних даних.

Ст. 8 ЄКПЛ охоплює захист широкого кола елементів нашого приватного життя, які можна згрупувати в три великі категорії, а саме: (i) (загальна) приватність особи, (ii) фізична, психологічна чи моральна цілісність особи та (iii) особистість і автономія людини [9].

Різні програми та використання ШІ можуть впливати на ці категорії, і на сьогоднішній день, їм приділено мало уваги.

Навіть якщо наші обличчя видаляються одразу після зйомки, технологія все одно втручається в нашу психологічну цілісність. І хоча для розпізнавання обличчя вплив на наше «загальне» право на конфіденційність і нашу психологічну цілісність може бути більш очевидним, можна стверджувати, що невибіркове відстеження всіх аспектів нашого життя (через нашу поведінку в Інтернеті, дані про місцезнаходження, наші дані від смарт-годинників, пристроїв відстеження стану здоров'я, розумних колонок, термостатів, автомобілів тощо) можуть мати такий самий вплив на наше право на конфіденційність, зокрема на нашу психологічну цілісність [105].

Інші форми біометричного розпізнавання, керованого ШІ, мають ще більший вплив на нашу психологічну цілісність. Розпізнавання мікровиразів, тону, голосу, частоти серцевих скорочень, температури тощо в даний час використовується для оцінки або навіть прогнозування нашої поведінки, психічного стану та емоцій.

Слід зазначити, що «Загальний регламент про захист даних» від 2016 року (GDPR) [10] обмежує обробку біометричних даних лише певною мірою [104]. Біометричні дані відповідно до GDPR – це персональні дані, отримані через спеціальну технічну обробку, що включає фізичні, фізіологічні або поведінкові ознаки, які дають змогу або підтверджують унікальну ідентифікацію особи.

Остання частина речення має вирішальне значення, тому що, якщо біометричне розпізнавання спрямоване не на ідентифікацію (а, наприклад, на категоризацію, профілювання чи розпізнавання), воно може не підпадати під визначення GDPR.

Повертаючись до конфіденційності даних, особисті та неособисті дані використовуються не лише для навчання систем Штучного інтелекту, але й для профілювання та оцінки людей для різних цілей, таких як прогнозування поліцейських, отримання страховки, соціальні виплати, прогнозування продуктивності при наймі та звільненні.

Крім того, величезна кількість «точок даних» про те, як ми займаємося повсякденним життям, використовується не лише для надсилання нам цільової реклами, але й для того, щоб підштовхнути/вплинути/спонукати/підштовхнути нас до певної інформації та, таким чином, впливати на наші варіанти, впливаючи на нашу моральність цілісність.

Саме в цей момент починає проявлятися вплив ШІ на свободу особи. Свобода особи відображена в Європейській конвенції з прав людини у різних правах, таких як свобода вираження поглядів (ст. 10) і свобода зібрань та асоціацій (ст. 11). Ст. 10 ЄКПЛ передбачає право на свободу вираження поглядів та інформації, включаючи свободу дотримуватись своїх поглядів, отримувати інформацію та ідеї [9]. Штучний інтелект, який використовується для профілювання, спостереження, відстеження та ідентифікації людей, перевірки, визначення, сортування та впливу на поведінку чи підштовхування до неї, не лише потенційно впливає на право на моральну цілісність, як описано вище, але також може негативно вплинути на ці конкретні свободи [104].

Використання розпізнавання обличчя в громадських місцях може заважати свободі думки та вираження особи просто через той факт, що захист «групової анонімності» більше не існує, якщо потенційно всіх у групі можна впізнати.

Що стосується права отримувати та передавати інформацію та ідеї, Штучний інтелект, який використовується в медіа та управлінні новинами, надає людям усе більше «персоналізованого» онлайн-контенту та новин, викликає занепокоєння [104].

Штучний інтелект також стає все більш здатним створювати медіаматеріали (відео, аудіо, зображення), які імітують вигляд і/або голос реальних людей (так звані «глибокі фейки»). Це дозволяє використовувати оманливі методи для різних цілей, що може призвести до поширення фейкових новин, дезінформації та пропаганди, що впливає на здатність людей формувати

та розвивати власні думки, отримувати та розповсюджувати інформацію, а отже, на нашу свободу вираження думок.

Одним із найбільш відомих впливів ШІ на права людини є вплив на заборону дискримінації та право на рівне ставлення.

Наприклад, ШІ рекрутингу Amazon надавав перевагу чоловікам, а не жінкам, оскільки його навчали на профілях успішних співробітників Amazon, якими виявилися чоловіки. Система Штучного інтелекту не просто відфільтровувала жінок, вона розглядала характеристики успішних співробітників, такі як типові формулювання і відфільтровувала резюме, які не відображали цих характеристик.

Попри наведені впливи варто розглянути, як ШІ впливається в соціально-економічні права [104].

Якщо роботодавці постійно контролюють працівників, вони можуть більше вагатися щодо організації.

Системи Штучного інтелекту, які оцінюють і прогнозують продуктивність працівників, можуть поставити під загрозу право на рівні можливості та рівне ставлення в питаннях зайнятості та професій без дискримінації за ознакою статі, особливо коли ці системи закріплюють упередження в даних або їхнього творця.

Існує ризик втрати необхідних навичок, коли все більше і більше роботи та рішень, які раніше виконувалися або приймалися людьми, переймаються ШІ-системами. Це може не тільки призвести до менш кваліфікованої робочої сили, але й збільшити ризик системного збою, де лише кілька людей здатні працювати з системами ШІ та реагувати на події, коли ці системи виходять з ладу.

Хоча невідомо, чи буде, і якщо так, скільки робочих місць буде втрачено або створено в результаті ШІ, у руйнівний період трансформації невідповідність між вразливою робочою силою та необхідними навичками може призвести до технологічного безробіття [104].

Варто відзначити вплив ШІ та демократію. ШІ має негативний вплив на демократію, зокрема, коли це стосується: соціального та політичного дискурсу, доступу до інформації та впливу на виборців, нерівності та сегрегації та системного збою або зриву [105].

Добре функціонуючі демократії вимагають добре поінформованого громадянства, відкритого соціального та політичного дискурсу та відсутності непрозорого впливу на виборців. В інформаційному суспільстві громадяни можуть обирати лише споживання невеликої кількості всієї доступної інформації. Пошукові системи, канали соціальних мереж, системи рекомендацій і багато сайтів новин використовують Штучний інтелект, щоб визначити, який контент створюється та показується користувачам (персоналізація інформації).

Персоналізація інформації на основі Штучного інтелекту забезпечується завдяки постійному моніторингу та профілюванню кожної людини [105].

Якщо особисті передбачення Штучного інтелекту стануть дуже потужними та ефективними, вони можуть навіть підірвати людську свободу волі та автономію, необхідні для прийняття значущих рішень виборцями.

Крім того, Штучний інтелект може підірвати справедливий виборчий процес. Політичні кампанії чи іноземні актори можуть використовувати персоналізовану рекламу для надсилання різних повідомлень окремим групам виборців без публічної підзвітності. Однак слід зазначити, що залишається невизначеним, наскільки точно є впливовою мікротаргетована реклама [105].

ШІ також можна використовувати для створення та поширення дезінформації та глибоких фейків у формі тексту, зображень, аудіо чи відео. Оскільки їх важко ідентифікувати громадянам, журналістам чи державним установам, вводити в оману та маніпулювати громадськістю стає легше, а рівень правдивості та довіри до ЗМІ та демократичного дискурсу може погіршитися.

Очікується, що нерівність і сегрегація ШІ покращить продуктивність економіки, проте ці прирости продуктивності розподілятимуться нерівномірно, причому більшість переваг отримають заможні верстви населення.

Подібним чином показано, що вибір даних і дизайну в поєднанні з недостатньою прозорістю алгоритмів чорної скриньки призводять до збереження несправедливого упередження щодо вже незахищених груп суспільства, таких як жінки та етнічні меншини [105].

ШІ може призвести до нерівності та сегрегації, а отже, загрожувати необхідному рівню економічної та соціальної рівності, необхідному для процвітаючої демократії.

Сьогодні варто оцінити вплив ШІ та верховенство права. Державні інституції дотримуються вищих стандартів, коли мова йде про їхню поведінку щодо окремих осіб і суспільства, що відображається в таких принципах, як виправданість, пропорційність і рівність.

ШІ може підвищити ефективність інституцій, але, з іншого боку, він також може підірвати процедурну легітимність і довіру до демократичних інституцій та авторитет закону [105].

Суди, правоохоронні органи та органи державного управління могли б стати більш ефективними, але ціною того, щоб вони були більш непрозорими та менш людськими органами, автономією та наглядом. Так само, якщо раніше суди були єдиними, хто визначав, що вважається незаконною мовою ненависті, сьогодні переважно приватні системи Штучного інтелекту визначають, чи припиняється мова платформами соціальних медіа. Ці системи Штучного інтелекту де-факто конкурують за владу із суддями та законом.

Загалом Штучний інтелект може сприяти розвитку судових систем, які працюють поза межами та захистом правління закон. Автоматизоване вирішення спорів онлайн, яке надають приватні компанії, регулюється умовами надання послуг, а не законом, який не надає споживачам однакових прав і процедурного захисту в державних судах [105].

Ще у 2018 р. Європейська комісія з питань ефективності правосуддя визначила п'ять принципів використання Штучного інтелекту в судовій системі в «Європейській етичній хартії [8] щодо використання ШІ в судових системах та їхньому середовищі» [106].

Основні принципи включають:

- Принцип дотримання основних прав людини при використанні ШІ.
- Принцип недискримінації, а саме запобігання розвитку будь-якої дискримінації між окремими особами чи групами осіб.
- Принцип якості та безпеки, який стосується обробки судових рішень і даних у безпечному технологічному середовищі.
- Принцип "під контролем користувача".
- Принцип прозорості, неупередженості та справедливості.

Експертна група високого рівня з ШІ також закликала державні органи дотримуватися семи вимог до надійного Штучного інтелекту при його розробці, закупівлі або використанні.

Дані вимоги включають:

1. Дотримання прав людини та етичних принципів. ШІ повинно розроблятися та застосовуватися з урахуванням прав людини, демократичних цінностей та етичних норм. Це означає усунення дискримінації та забезпечення рівності прийнятих рішень. Такий підхід завдяки створенню довіри до ШІ, але його впровадження залежить від доступності якісних, непередбачених даних.

2. Технічна надійність та безпека. Системи ШІ повинні бути стійкими до технічних збоїв і захищеними від зовнішніх загроз. Зменшує ризик шкідливих наслідків для користувачів і суспільства. Одночасне забезпечення цих вимог може ускладнювати розробку через необхідність багаторівневого тестування.

3. Конфіденційність і управління даними. ШІ має гарантувати конфіденційність персональних даних, забезпечуючи контроль користувачів над інформацією. Це включає дотримання стандартів захисту даних, таких як

GDPR. Однак збалансування конфіденційності з функціональністю ШІ може бути технічно складним завданням.

4. Прозорість. Алгоритми ШІ повинні бути зрозумілими та пояснюваними. Користувачі мають право знати, як працює система і які фактори впливають на прийняття рішень. Протест, складність сучасних моделей, як-от нейронні мережі, робить пояснюваність значним викликом.

5. Підзвітність. Відповідальність для наслідків використання ШІ має бути чітко визначена. Це оплачується як розробників, так і організацій, що впроваджують ці технології. Така вимога важлива для запобігання зловживання, але потребує розробки юридичних механізмів, які враховують специфіку ШІ.

6. Соціальна і екологічна торгівля. ШІ повинна сприяти суспільному благу та екологічній стабільності. Його застосування має підтримувати економічний розвиток.

Однак Штучний інтелект може не тільки загрожувати верховенству права, але й зміцнити його. Якщо його розробити та використовувати відповідально, він може надати агентствам можливість виявляти корупцію разом із державними структурами. Подібним чином ШІ можна використовувати для виявлення та захисту від кібератак [106].

Вплив Штучного інтелекту на права людини, демократію та верховенство права останнім часом приділяється більше уваги, особливо в останньому Whitepaper щодо Штучного інтелекту Європейської комісії. Проте, як усунути цей вплив, на сьогоднішній день залишається невизначеною територією.

Саме тому сформовані можливі стратегії, яких можна дотримуватися в існуючих рамках прав людини, демократії та верховенства права [див. Додаток 3, с.124].

Згідно першої стратегії багато розробників Штучного інтелекту, розробників і користувачів (державних і приватних), здається, не знають про (потенційний) вплив Штучного інтелекту на права людини. Як перший крок, ітерація або (повторне) артикуляція, під час якої існуючі права людини ЄКПЛ

«перекладаються» в контекст Штучного інтелекту, є дуже корисною та може бути виконана за допомогою Рамкової конвенції 2016 р.

Щоб ефективно впоратися з впливом ШІ на права людини, демократію та верховенство права, можна продовжити розвиток існуючих механізмів дотримання, підзвітності та відшкодування, а також впровадити нові. Однак важливо врахувати, що використання Штучного інтелекту часто залишається прихованим або невідомим, що ускладнює або робить неможливим визначення його впливу на права людини, демократію та верховенство права. Тому заходи щодо відповідності, підзвітності та відшкодування повинні починатися з вимоги прозорості у використанні систем ШІ, які можуть впливати на ці сфери. Це включає створення реєстру ШІ, що дозволяє визначити рівень ризику та відповідний ступінь прозорості та підзвітності для кожної програми.

Оскільки Штучний інтелект має широкий спектр застосувань, таких як стеження, ідентифікація, профілювання, підштовхування і прийняття рішень, засоби правового захисту повинні бути адаптовані до різних програм [82].

Належні правові заходи мають включати припинення протиправної поведінки і гарантії неповторення, наприклад, через обов'язок держав прийняти та впровадити законодавство, яке захищає права людини від майбутнього впливу ШІ. Має бути забезпечене відшкодування шкоди, завданої порушеннями, як для окремих осіб, так і для громад. Для деяких програм Штучного інтелекту може бути недостатньо просто ефективного правового захисту.

За третьою стратегією, щоб запобігти системним збоям або збоям через централізацію керованих ШІ процесів прийняття рішень у життєво важливих структурах, слід запровадити розподілені процеси прийняття рішень, а не централізовані, щоб запобігти ризику катастрофічного збою. У ці процеси повинні бути вбудовані відповідні структури людського нагляду. Людський нагляд допомагає гарантувати, що система Штучного інтелекту не підриває автономію людини та не спричиняє інших несприятливих наслідків. Щоб

подолати ризик нерівності, уряди повинні активно припиняти використання програм ШІ, які посилюють нерівність. Запобігання впливу на вибори чи публічному маніпулюванню через персоналізовану інформацію, керовану Штучним інтелектом, є непростим завданням [46].

Отже, демократичні процеси можуть розвиватися від поширення дезінформації та маніпуляції громадською думкою за допомогою ШІ, включаючи вплив на виборчі кампанії. З іншого боку, прозоре та підзвітне впровадження ШІ посилення зміцнення демократичних інституцій. Європейська конвенція з прав людини підкреслює важливість захисту приватності, рівності та людської гідності, на яку ШІ може мати як позитивний, так і негативний вплив.

ШІ може призвести до економічної нерівності, створюючи значні вигоди для окремих груп, але посилюючи вразливість інших. У демократичному суспільстві це створює загрозу соціальної рівності та злагоди. Для мінімізації ризиків важливо інтегрувати розподілені процеси прийняття рішень та механізми людського контролю.

Крім того, ШІ може зміцнити верховенство права, якщо використовувати його для боротьби з корупцією або для автоматизації процесів розслідування. Однак централізовані системи прийняття рішень можуть підірвати прозорість та легітимність демократичних інституцій. Для уникнення цього необхідно забезпечити прозорість і публічну підзвітність системи ШІ.

Баланс між можливостями та ризиками ШІ вимагає впровадження етичних принципів, технічної надійності, конфіденційності даних і відповідності. Міжнародні ініціативи, такі як європейська Етична хартія [8] щодо ШІ, почали цікавити орієнтири розвитку технологій у відповідності до прав людини. Узгоджена стратегія, що включає прозорість, нагляд і запобігання дискримінації, є основою для інтеграції ШІ в демократичне суспільство. Водночас важливо розробляти нові механізми правового захисту, враховуючи складність впливу ШІ на різні сфери.

2.3. Вплив ШІ на захист інтелектуальної власності та авторських прав

У царині проривних технологій Штучний інтелект (ШІ) стоїть передвісником інновацій і трансформації. Потенціал ШІ, що революціонізує традиційні практики та моделі в усіх секторах, став незаперечним. Оскільки Штучний інтелект продовжує розвиватися безпрецедентними темпами, він має глибокий вплив на захист інтелектуальної власності (ІВ).

Штучний інтелект — це рішення, яке можна реалізувати як інструмент для багатьох цілей, включаючи створення нового та «оригінального» контенту, готового до використання, модифікації або переробки дизайнерами, комунікаційними та маркетинговими командами [57]. Дійсно, для багатьох 2023 р. став роком генеративного Штучного інтелекту (Generative AI). Generative AI здатний створювати нові типи «оригінальних» творів, таких як тексти, зображення та відео, дотримуючись підказок, введених користувачем вручну (інструкцій або запитів, які користувач надає моделі ШІ для отримання бажаного результату). Ці твори не є спонтанними, за лаштунками є програмне забезпечення, яке реалізує алгоритм, попередньо навчений на величезних обсягах даних. Завдяки цьому навчанню генеративний Штучний інтелект може генерувати прогнози та зв'язувати між введеним текстом (підказками) від користувачів і всіма наявними раніше знаннями. Завдяки цій асоціації ШІ забезпечує вихід/«оригінал» твору відповідно до наданих інструкцій. Чим кращий інструмент Штучного інтелекту, тим краще результат відповідає вимогам, включеним до інструкцій, наданих користувачем.

Штучний інтелект стає дедалі більше у публічній сфері, і його роль у бізнесі лише зростатиме. За оцінками уряду Великої Британії, Штучний інтелект привніс 3,7 мільярда фунтів стерлінгів в економіку Великої Британії минулого року – враховуючи це, зацікавленість і, отже, фінансування ШІ продовжуватиметься швидко [75].

Ці інноваційні розробки позитивно вплинуть на те, як ми взаємодіємо з технологіями та Інтернетом. Основні переваги використання ШІ в інтелектуальній власності та авторському праві включають:

1. Автоматичний пошук IP. Алгоритми на основі Штучного інтелекту можуть проводити комплексний пошук відомостей про рівень техніки та існуючих баз даних IP ефективніше й точніше, ніж люди-дослідники. Це дає змогу швидше виявляти потенційні конфлікти, покращує належну обачність і допомагає у визначенні проблем патентоспроможності чи порушення.

2. Аналіз рівня техніки. ШІ може допомогти в аналізі величезних обсягів технічної інформації та документів для виявлення відповідного рівня техніки. Це може допомогти патентним експертам, повіреним і винахідникам в оцінці новизни та неочевидності винаходів, які є ключовими аспектами патентоспроможності.

3. Автоматизоване складання патентів. ШІ може генерувати проекти патентних заявок, аналізуючи великі набори даних існуючих патентів і юридичних документів. Хоча адвокати продовжують відігравати ключову роль у вдосконаленні та завершенні патентних заявок, інструменти Штучного інтелекту можуть підвищити ефективність процесу, зменшити кількість помилок та надати рекомендації для створення більш обґрунтованих патентних претензій.

4. Виявлення порушення авторських прав. Алгоритми ШІ можуть допомогти виявити випадки порушення авторських прав на цифрових платформах шляхом порівняння вмісту з базою даних творів, захищених авторським правом. Ця технологія дозволяє власникам авторських прав контролювати їхні роботи та вживати відповідних заходів.

5. Очищення та моніторинг торгових марок. Штучний інтелект може допомогти в пошуку дозволів на торговельні марки, аналізуючи величезні бази даних торгових марок для виявлення потенційних конфліктів. Алгоритми Штучного інтелекту також можуть контролювати онлайн-платформи, веб-сайти

та соціальні мережі на предмет несанкціонованого використання торговельних марок, забезпечуючи проактивне правозастосування.

6. Аналіз контракту. Інструменти на основі Штучного інтелекту можуть оптимізувати аналіз контрактів шляхом автоматичного вилучення та класифікації відповідної інформації, включно з пунктами, пов'язаними з інтелектуальною власністю. Це економить час і дозволяє адвокатам зосередитися на завданнях вищого рівня, таких як переговори та складання контрактів.

7. Підтримка судових спорів щодо інтелектуальної власності. ШІ може допомогти в судових процесах щодо ІВ, аналізуючи великі обсяги правових текстів, судових рішень і прецедентів. Він може допомогти адвокатам у юридичних дослідженнях, розробці стратегії справи та прогнозуванні результатів на основі історичних даних [65].

Тривають ключові дебати щодо впливу Штучного інтелекту на існуючі системи інтелектуальної власності (ІВ), а також щодо того, чи можна запатентувати роботу, створену ШІ, або захистити авторським правом так само, як і роботу, створену людиною.

Авторські права призначені для захисту оригінальних творчих творів, зафіксованих у матеріальній формі, таких як літературні, музичні та художні твори. Автори цих оригінальних творів мають право відтворювати, розповсюджувати та публічно демонструвати свої твори протягом певного часу. Завдяки ШІ межа між авторством людини та машини стає розмитою. Системи ШІ можуть створювати твори мистецтва, музику, літературу та інші творчі результати.

Однак важливо пам'ятати, що існують також значні потенційні ризики, оскільки прогрес рухається такими швидкими темпами, що більшість інструментів Штучного інтелекту на ринку ще не повністю вивчені та майже не мають регулювання. Ілон Маск (Elon Musk) і Стівен Гокінг (Stephen Hawking)

є одними з тих, хто закликав до посилення досліджень і регулювання, щоб забезпечити розробку ШІ в етичній манері.

У 2021 р. Європейська комісія (ЄК) запропонувала Закон про Штучний інтелект, розроблений для того, щоб заохотити розробників Штучного інтелекту в Європі піклуватися про прозорість і безпеку користувачів, відносячи додатки Штучного інтелекту до трьох категорій ризику (неприйнятні, високоризикові та невисокі) [84]. Однак, у цьому Законі не згадується вплив на інтелектуальну власність, що залишає невизначеним щодо того, як буде захищена інтелектуальна власність власника цих прав.

У квітні 2023 р. Кіберадміністрація Китаю оприлюднила для громадського обговорення проект заходів щодо адміністрування послуг генеративного Штучного інтелекту. Проект заходів намагається збалансувати розвиток технології зі збереженням контролю над нею та гарантування громадського порядку. Одна з ключових вимог, яку необхідно виконати, полягає в тому, щоб компанії, які створюють генеративну технологію ШІ, запроваджували заходи безпеки для забезпечення відповідності – не лише для запобігання створенню контенту, який не суперечить державній політиці, але й з урахуванням наслідків для захисту інтелектуальної власності [49].

Інструмент інтелектуального аналізу авторських прав (AI), розроблений у Китаї, є значним прогресом у законодавстві про інтелектуальну власність (ІВ). Використовуючи Штучний інтелект, ця інноваційна система допомагає перевіряти та захищати твори, захищені авторським правом. Він забезпечує надійне цифрове рішення проблем, з якими стикаються суди та юристи, спрощуючи процес оцінки оригінальності, інноваційності та потенційного порушення авторських прав.

Ключові характеристики цього Інструменту аналізу авторських прав:

➤ Розширене розпізнавання зображень: Інструмент використовує розширене розпізнавання зображень для пошуку «зображення до зображення»,

відстеження походження візуальних творів і ідентифікації оригінальних творців, що є важливим для встановлення права власності.

➤ Оцінка інноваційності та подібності: оцінюється ступінь інноваційності у творі та кількісно оцінюється подібність між зображеннями, допомагаючи визначити потенційне порушення авторських прав.

➤ Безпроблемна інтеграція для обґрунтованих суджень: інтегрований у платформи управління судами, інструмент надає суддям аналітичні ресурси під час судових розглядів, покращуючи процес прийняття рішень і оптимізуючи вирішення справ з високим рівнем успіху у виявленні подібних проєктів.

Переваги:

➤ Простота використання та економічна ефективність: зручний інтерфейс дозволяє суддям і працівникам суду легко орієнтуватися, зменшуючи потребу в тривалому навчанні та знижуючи судові витрати за рахунок спрощення процесу збору доказів.

➤ Ефективність і точність аналізу авторських прав. Завдяки автоматизації аналізу випадків авторського права інструмент значно скорочує час, необхідний для оцінки, і мінімізує людські помилки, що забезпечує точніші рішення

➤ Посилення захисту авторських прав: Інструмент посилює захист авторських прав, надаючи право професіоналам у сфері права захищати права творців і ефективно боротися з порушеннями [82].

Є два способи, якими законодавство про авторське право може розглядати твори, де людська взаємодія мінімальна або відсутня. Він може або відмовити в захисті авторських прав для творів, створених комп'ютером, або приписати авторство таких творів творцю програми.

Наразі закони про авторське право зазвичай вимагають авторства людини, залишаючи творіння, створені ШІ, у сірій юридичній зоні. Бюро авторських прав Сполучених Штатів надало вказівки, згідно з якими не люди не вважаються авторами, які можуть отримати захист авторських прав. Однак у

цьому керівництві враховується, що ШІ може допомогти у створенні роботи, якщо він допомагає людині, яка отримує «традиційні елементи авторства».

Ця сіра зона авторського права між Штучним інтелектом і авторством людини викликає занепокоєння Бюро авторських прав США. Кількість заявок на реєстрацію створених Штучним інтелектом робіт, ймовірно, зросте, оскільки доступ до генеративного Штучного інтелекту збільшиться [33].

У «Звіті Бюро авторських прав США» з 2024 р. у висновках відзначається нагальна потреба в новому захисті авторських прав на федеральному рівні. Широка доступність генеративних інструментів Штучного інтелекту, які спрощують створення цифрових копій зображень і голосів людей, окреслила прогалини в існуючому законодавстві та викликала занепокоєння щодо шкоди, яку може завдати їх несанкціоноване використання [46].

Продуктивність генеративного Штучного інтелекту залежить як від вхідних даних користувача, так і від величезного набору даних, більшість із яких зібрано з Інтернету, які використовуються для їх створення. Використання процесів і правил для збору інформації з великого набору даних тексту, створеного людиною, означає, що створений контент міститиме деякі елементи, ті ж елементи, що й уже існуюча література чи мистецтво.

Технічно це означає, що початковий власник може подати позов до генератора ШІ або творця ШІ за порушення авторських прав. Однак досі незрозуміло, як розглядатиметься ця судова справа та хто нести відповідальність. Цілком ймовірно, що розробникам інструментів Штучного інтелекту потрібно буде продемонструвати, що впроваджено належні засоби захисту, щоб запобігти порушенням прав інтелектуальної власності їхніми системами. Розробники повинні забезпечити встановлення відповідних правил та укладання контрактів з третіми сторонами, такими як художники, власники бібліотек зображень та баз даних, для використання даних, що використовуються для створення результатів. Користувачі таких систем

повинні перевіряти, що будь-який контент, який вони створюють, не порушує права третіх осіб, особливо в комерційних цілях.

Твір має бути оригінальним витвором автора, щоб охоронятися авторським правом. Це ставить питання, чи можна вважати твір, створений ШІ автентичним? Незважаючи на те, що це прямо не заборонено, поточна судова практика категорично проти цього. Як Європейський суд Європейського Союзу, так і Верховний суд США вважають, що захист авторських прав поширюється лише на оригінальні твори з певним ступенем інтелектуальної творчості, творчих здібностей розуму та відображення особистості автора (у справах *Inforaq* і *Feist Publications*, відповідно). Коли Штучний інтелект є єдиним творцем тоді елемент оригінальності відсутній, а це означає, що автор також відсутній [87].

Виходячи з цього, на даний момент існує два можливих рішення цієї проблеми. По-перше, немає автора. Твір, створений ШІ, вважатиметься суспільним надбанням. Крім того, робота буде вважатися похідною від усіх матеріалів ШІ, які використовуються для навчання. Однак існує очевидна проблема в тому, щоб довести, що щось було створено ШІ. Розпізнати, чи було щось створено Штучним інтелектом чи ні, майже неможливо. Це ускладнює проблему ще більше, оскільки навіть надання ШІ одних і тих самих підказок може призвести до інших результатів на основі попередніх підказок, наданих ШІ. Навіть програмне забезпечення для виявлення ШІ може не впоратися з цим завданням, особливо якщо код програми недоступний для широкої громадськості. Один зі способів уникнути цієї проблеми — запропонувати художникам записувати весь процес створення. Це, очевидно, вкрай непрактично, але поки що це єдиний спосіб уникнути будь-яких сумнівів щодо походження творів.

За останні кілька років було багато прикладів складних відносин між ШІ та інтелектуальної власністю. Один приклад впливає з того періоду, коли доктор Стівен Талер намагався вказати Штучний інтелект під назвою

«DABUS» як винахідника в патентній заявці. Після того, як Бюро патентів і торгових марок Сполучених Штатів відхилило заявку через відсутність у списку людини-винахідника, Талер оскаржив це рішення та закон, що вимагає наявності людини-винахідника. Апеляційний суд Федерального округу США дійшов висновку, що Закон про патенти чіткий: «винахідники повинні бути фізичними особами» [33].

Інший більш відомий приклад походить із музичної індустрії. У квітні 2023 р. Дрейк випустив нову пісню за участю The Weeknd під назвою «Серце на рукаві» («Heart on My Sleeve»). Проблема в тому, що ані Дрейк, ані The Weeknd не продюсували трек. Запис був створений Штучним інтелектом за допомогою технології імітації голосу, а потім пісня була завантажена на кілька музичних потокових сайтів. Серйозність цього прикладу неможливо переоцінити. Дрейк і The Weeknd отримали кілька нагород Греммі, а їхня музика збирає мільярди завантажень. Із зростанням популярності «Heart on My Sleeve» та інших, створених таким же чином, цей тип порушень навряд чи припиниться [69].

Щоб вирішити складні проблеми інтелектуальної власності, пов'язані з Штучним інтелектом, політики та юридичні експерти повинні розглянути кілька потенційних змін політики :

1. Уточнення авторства: внесіть зміни до законів про авторське право, щоб чітко визнавати ШІ потенційним автором, дозволяючи творам, створеним ШІ, мати право на захист авторських прав чи ні. Це може передбачати створення нової категорії авторства, яка враховує участь ШІ.

- 2 . Розвиток законів про торговельні марки: оновіть закони про торговельні марки, щоб вони охоплювали вміст, створений Штучним інтелектом, і створили механізми моніторингу та забезпечення дотримання прав на товарні знаки в матеріалах, створених ШІ, забезпечуючи захист власників брендів.

3. Перевизначення права власності на патент: запровадьте законодавчі акти чи вказівки для роз'яснення права власності на патент, коли задіяний Штучний інтелект. Розглянемо гібридний підхід, який визнає внесок як систем ШІ, так і винахідників, забезпечуючи справедливий розподіл прав.

4. Стандартизація атрибуції Штучного інтелекту: розробка галузевих стандартів і найкращих практик для приписування створених Штучним інтелектом творів їхнім творцям. Це може допомогти встановити прозорість і підзвітність у використанні систем ШІ, одночасно пом'якшуючи судові спори.

5. Сприяння співпраці: заохочуйте співпрацю між юридичними експертами, дослідниками Штучного інтелекту, політиками та зацікавленими сторонами галузі, щоб орієнтуватися в швидкому розвитку Штучного інтелекту та інтелектуальної власності. Ця співпраця повинна визначити етичні принципи та забезпечити, щоб регулювання інтелектуальної власності відповідало технологічному прогресу. В умовах, коли Штучний інтелект, ймовірно, залишиться постійною складовою повсякденного життя, важливо створити план і підтримувати діалог між розробниками ШІ та законодавцями. Це особливо актуально на сучасному етапі розвитку технологій, адже ШІ навряд зникне.

Другий варіант, передання авторства програмісту, представлений у таких країнах як Гонконг (SAR), Індія, Ірландія, Нова Зеландія та Великобританія [66]. Британія є однією з небагатьох країн у світі, яка має спеціальну правову базу щодо авторського права на твори, створені комп'ютерами. Коли цей механізм було запроваджено в Законі «Про авторське право, зразки та патенти» 1988 р. (CDPA) [67], оновлення якого чинності 29 жовтня 2024 р., його проголошували «першим у світі законодавством про авторське право, яке намагається впоратися з появою Штучного інтелекту» [47]. Цей підхід найкраще втілено в законі Великобританії про авторське право, розділ 9 (3) Закону «Про авторське право, зразки та патенти» (CDPA) від 1988 р., що оновлений з усіма змінами, які набули чинності 29 жовтня 2024 р. зазначає, що

«у випадку літературного, драматичного, музичного чи художнього твору, створеного за допомогою комп'ютера, автором вважається особа, яка вжила заходів, необхідних для створення твору» [66].

Надання авторських прав людині, яка уможливила роботу ШІ, здається найбільш розумним підходом, причому модель Великобританії виглядає найефективнішою. Такий підхід забезпечить бажання продовжувати інвестувати в технологію, та впевненість, що вони отримають прибуток від своїх інвестицій.

Управління інтелектуальної власності Великобританії (UK Intellectual Property Office UKIPO) повідомило, що все ще має намір завершити роботу над новим кодексом авторських прав на Штучний інтелект. Робоча група, до складу якої входять представники технологічного, творчого та дослідницького секторів, створена UKIPO, веде переговори щодо нового добровільного кодексу авторського права на ШІ, який, як сподівається уряд, забезпечить незаконодавче вирішення питання про те, як найкраще збалансувати потребу розробників ШІ в доступі до якісних даних, на основі яких можна навчати системи Штучного інтелекту, з потребою творців контенту отримувати належну винагороду за використання їх авторських робіт.

Ще в березні 2023 р. сер Патрік Валланс (Sir Patrick Vallance) опублікував свій огляд про інноваційне регулювання цифрових технологій. У ньому він визначив нагальну потребу в пошуку практичних рішень для бар'єрів, з якими стикаються фірми, що займаються Штучним інтелектом, під час доступу до матеріалів, захищених авторським правом і правами на базу даних, і рекомендував скласти з цією метою кодекс практики [111].

У Великій Британії чинне законодавство не дозволяє несанкціоноване копіювання захищеного авторським правом вмісту для навчання моделей Штучного інтелекту, за винятком випадків, коли це робиться в суто некомерційних цілях. Це означає, що якщо комерційна компанія ШІ бажає скопіювати матеріали третьої сторони, які є загальнодоступними в Інтернеті,

для навчання своїх моделей, їй може знадобитися отримати дозвіл на це від кожного відповідного правовласника. Однак, оскільки відповідного власника прав може бути важко ідентифікувати, і оскільки для цієї мети немає впорядкованої системи ліцензування, а також оскільки системи машинного навчання потребують величезної кількості даних для навчання, цей процес може бути важким і громіздким на практиці.

В інших країнах законодавство більш гнучке. Наприклад, в ЄС існує виключення, яке дозволяє копіювання з цією метою, за винятком випадків, коли власник вмісту відмовився від цього. У США за деяких обставин можуть застосовуватися ширші винятки, але це залишається невизначеним і є предметом серйозних судових спорів між творцями та технологічними компаніями [37].

Раніше урядові пропозиції щодо внесення змін до законодавства, які дозволять ширше використовувати твори, захищені авторським правом, для навчання Штучного інтелекту зустріли негативну реакцію з боку креативного сектору. Однак остання реакція уряду «дає зрозуміти, що його поточний підхід до цього питання повністю поважатиме права творців контенту, незважаючи на те, що розробка технологій Штучного інтелекту залишається пріоритетом» відзначив Гілл Денніс (Gill Dennis), експерт із права інтелектуальної власності [50]. Публікація Кодексу правил спочатку мала бути опублікована влітку 2023 р., а потім була перенесена на початок 2024 р., що відображає складність досягнення консенсусу з даного питання. Практичний кодекс має на меті зробити ліцензії на аналіз даних більш доступними. Це допоможе подолати перешкоди, з якими зараз стикаються фірми та користувачі ШІ, і забезпечить захист правовласників. Це гарантує, що система авторського права Великобританії заохочує та винагороджує інвестиції у творчість. Він також підтримує прагнення Великої Британії стати світовим лідером у галузі досліджень та інновацій ШІ.

UKIPO не вдалося укласти угоду між зацікавленими сторонами креативної індустрії та компаніями Штучного інтелекту щодо використання контенту, захищеного авторським правом, для навчання моделей ШІ на початку 2024 р. Уряд Великобританії підтвердив, що, хоча Робоча група забезпечила цінний форум для зацікавлених сторін для обміну думками, однак вона не змогла погодити ефективний добровільний кодекс. Ця подія також знаменує невдачу в амбіціях уряду стати «глобальною супердержавою» і взяти на себе провідну роль у розробці узгодженої структури для використання ШІ в усьому світі.

Отже, оскільки Штучний інтелект продовжує трансформувати галузі та розширювати можливості людської творчості, необхідно розв'язати складні питання, що виникають на перетині Штучного інтелекту та інтелектуальної власності. Оновлення чинного законодавства про інтелектуальну власність, уточнення авторства, розвиток положень щодо торговельних марок, перегляд прав на патенти та сприяння співпраці можуть допомогти створити правову основу, яка стимулює інновації та одночасно захищає права творців і суспільства. Баланс інтересів людей і машин в інтелектуальній власності має вирішальне значення для справедливого та стійкого розвитку технології ШІ.

Однак, оскільки технологія ШІ розвивається так швидко, що законодавство, яке регулює Штучний інтелект повинно постійно оновлюватися, щоб залишатися актуальним і корисним. Надмірно директивні правила гальмуватимуть розвиток, тоді як їхня відсутність може призвести до швидкого зриву та непередбачуваних негативних наслідків.

РОЗДІЛ 3 ПРАВОВІ АСПЕКТИ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ

3.1. Тенденції розвитку зарубіжного та вітчизняного законодавства щодо регулювання систем Штучного інтелекту

Системи штучного інтелекту (ШІ) стають дедалі важливою частиною сучасного суспільства, впливаючи на економіку, право, етику та технології. У зв'язку з цим регулюванням ШІ набуває стратегічного значення як у зарубіжних, так і в вітчизняних правових системах. Світова практика демонструє, що уряди та міжнародні організації активно розробляють законодавчі акти, спрямовані на забезпечення прозорості, відповідальності та безпечного використання ШІ. Це включає питання захисту персональних даних, алгоритмічної дискримінації та регулювання автономних рішень, які можуть впливати на права людини.

Водночас в Україні розвиток законодавства щодо ШІ переходить на ранніх етапах. Проте помітний інтерес до адаптації найкращих світових практик з урахуванням національних особливостей. Серед ключових завдань для вітчизняного регулювання — визначення стандартів для етичного використання ШІ, інтеграція його в державне управління та бізнес, а також забезпечення конкурентоспроможності країни в глобальному технологічному середовищі. Цей процес потребує співпраці між урядом, науковцями, бізнесом та громадянським суспільством.

Підходи до управління ризиками Штучного інтелекту (Artificial Intelligence AI), сформовані новим законодавством, регулятивним наглядом, цивільною відповідальністю, м'яким правом і галузевими стандартами, стають ключовими аспектами міжнародної дипломатії та торгової політики. На додаток до заохочення інтегрованих ринків технологій, більш уніфікований міжнародний підхід до управління Штучним інтелектом може посилити

регулятивний нагляд, спрямовувати дослідження на спільні виклики, сприяти обміну найкращими практиками та забезпечити взаємодію інструментів для надійної розробки Штучного інтелекту.

Особливо впливові в цьому ландшафті є ЄС і США, які зараз впроваджують основоположні політики, які створюють прецеденти для майбутнього управління ризиками ШІ на їхніх територіях і в усьому світі.

Підходи до управління стосуються широкого спектру використання Штучного інтелекту з міжнародними наслідками, включаючи більш складний Штучний інтелект у споживчих продуктах; поширення ШІ в регульованих соціально-економічних рішеннях; розширення ШІ на різноманітних онлайн-платформах; і загальнодоступні веб-системи ШІ, такі як генеративний ШІ та базові моделі [99].

30 травня 2022 року близько 350 фахівців у галузі Штучного інтелекту (ШІ) підписали лист, в якому висловили серйозну стурбованість щодо ризиків, пов'язаних з ШІ. У листі підкреслюється, що «зменшення ризику, пов'язаного з ШІ, має стати глобальним пріоритетом на рівні з іншими загрозами суспільного масштабу, такими як пандемії та ядерна війна» [26]. Цей лист є останнім у серії попереджень про потенційні ризики — як малоймовірні, так і реальні — що можуть виникнути внаслідок розробки та впровадження ШІ. Незалежно від того, чи здійсняться ці побоювання, існує загальний консенсус серед ключових гравців у приватному і державному секторах про необхідність вже зараз розробити регуляцію ШІ. Проте підходи до відповідального управління ризиками ШІ та відповідні правила вже відрізняються в різних юрисдикціях.

США інвестували в нерегулятивну інфраструктуру, таку як нова структура управління ризиками Штучного інтелекту, оцінка програмного забезпечення для розпізнавання обличч і значне фінансування досліджень ШІ.

Сполучені Штати просувають національні стандарти Штучного інтелекту через виконавчу владу, але також прийняли деякі правила щодо ШІ на рівні штату (як через всеосяжне законодавство про конфіденційність, так і для

конкретних випадків використання, пов'язаних зі ШІ). Підхід федерального уряду США до управління Штучного інтелекту є галузевий і широко розподілений між федеральними агентствами. У цьому підході є переваги, однак він також сприяє нерівномірному розвитку політики ШІ. Хоча існує кілька керівних федеральних документів Білого дому щодо шкоди ШІ, вони не створили рівномірного чи послідовного федерального підходу до ризиків ШІ: «Загалом федеральні агентства досі не розробили необхідних планів регулювання ШІ» [54].

На сьогоднішній день існує три ключові виконавчі розпорядження (ЕО), які впливають на Штучний інтелект у федеральному уряді та його використання та розвиток приватними компаніями. Перше з них датується 2019 р., друге — 2020 р., а останнє, більш помітне, — 2023 р. Виконавчий наказ від лютого 2019 р. «Збереження американського лідерства у сфері Штучного інтелекту» (Maintaining American Leadership in Artificial Intelligence EO 13859). Його метою є встановлення принципів і стратегій для посилення можливостей США у сфері ШІ для сприяння науковим відкриттям, економічній конкурентоспроможності та національній безпеці [83]. Американська ініціатива Штучного інтелекту має керуватися п'ятьма принципами: сприяти технологічним проривам у галузі ШІ в уряді; стимулювати розвиток технічних стандартів; США повинні навчити свою робочу силу навичкам, пов'язаним із ШІ; сприяння суспільній довірі та довірі до технологій Штучного інтелекту шляхом захисту свобод і конфіденційності; відкрити ринки для американських компаній ШІ.

Оскільки Штучний інтелект (ШІ) набув широкого поширення в суспільстві, і виконавча гілка влади, і Конгресс почали вживати заходів щодо впровадження ШІ в урядових операціях. Прийнятий як частина Закону Вільяма М. (Мака) Торнберрі про санкціонування національної оборони на 2021 фінансовий рік (PL 116-283), Конгресс окреслив план «Національної ініціативи Штучного інтелекту, заохочуючи міжвідомчу співпрацю для реалізації більш

конкретної стратегії ШІ». Розділ Е, «Закон про національну ініціативу щодо Штучного інтелекту 2020 року» (The National Artificial Intelligence Initiative Act of 2020” (NAIIA)) містить законодавчі вказівки щодо наступних кроків, вимагає звітів для підзвітності та підкреслює необхідність включення відкритих даних і найкращих практик обміну даними в стратегію ШІ [85].

Також, прийнято Закон «Про ШІ в уряді» від 25 вересня 2020 р. (H.R.2575 - AI in Government Act of 2020). Цей акт створив Центр передового досвіду Штучного інтелекту (AI Center of Excellence within) в рамках Адміністрації загального обслуговування (the General Services Administration) та наказав Адміністративно-бюджетному управлінню (Office of Management and Budget OMB) випустити меморандум, в якому інформуватиме федеральні агентства про політику придбання та застосування Штучного інтелекту та визначати найкращі практики для зменшення ризиків. Закон має на меті реалізувати наступні завдання:

Допомога у впровадженні технологій ШІ.

Розвиток можливостей та використання AI.

Підвищення продуктивності та ефективності на благо громадян.

З огляду на те, що Штучний інтелект стане важливою майбутньою ініціативою, керівники ключових американських агентств, установ повинні стратегічно спланувати свої наступні кроки для інтеграції ШІ. Закон і запропонована політика окреслюють ключові нові ролі, обов'язки, звітність і вимоги, які агентствам необхідно буде враховувати під час впровадження бізнес-процесів ШІ [68].

3 грудня 2020 р. президент Трамп видав указ 13960 «Про сприяння використанню надійного Штучного інтелекту у федеральному уряді» (Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government). Указ встановлює принципи використання ШІ у федеральному уряді, встановлює загальну політику щодо реалізації принципів, вимагає від агентств створити перелік випадків використання Штучного

інтелекту та закликає Адміністрацію загального обслуговування та Управління персоналом посилити Експертизу впровадження ШІ в агентствах [89].

Ці документи також закликали агентства розглянути ключові аспекти зниження ризиків Штучного інтелекту за допомогою регулятивних і нерегуляторних заходів. Це включає використання наукових доказів для визначення можливостей Штучного інтелекту, забезпечення дотримання законів про недискримінацію, розгляд вимог щодо розкриття інформації та сприяння безпечній розробці та розгортанню Штучного інтелекту. Хоча ці документи відображали мінімалістичну регуляторну точку зору адміністрації Трампа, вони також вимагали від агентств розробити плани щодо регулювання додатків ШІ.

Загалом, федеральні агентства досі не розробили необхідних планів регулювання ШІ. У грудні 2022 року Інститут Штучного інтелекту, орієнтованого на людину, що був утворений при Стендфордському університеті 2019 р. [78] (Stanford University's Center for Human-Centered AI) опублікував звіт, у якому зазначено, що реалізація американської стратегії щодо інновацій ШІ, затверджених в Указах Президента зіткнулася з серйозними труднощами, і на сьогоднішній день не виконана. Зокрема, Офіс управління персоналом (The Office of Personnel Management OPM) не встановив серію професій ШІ або оцінку потреб у робочій силі, пов'язаної з ШІ, як того вимагає Закон «Про ШІ в уряді».

Адміністративно-бюджетне управління (The Office of Management and Budget OMB) ще не опублікувало меморандум, який вимагається відповідно до Закону для надання вказівок і принципів федерального придбання та використання ШІ, пом'якшення дискримінаційного впливу чи упередженості та сприяння інновацій в галузі ШІ. ОМВ також не опублікував дорожню карту, включаючи графік і графік залучення громадськості, для політичних вказівок, які мають бути переглянуті або створені для підтримки використання ШІ, як того вимагає ШІ в Урядовому розпорядженні.

88 % агенцій, які, ймовірно, підпадають під дію вимоги щодо подання планів залучення та впровадження ШІ, не зробили цього. Лише 5 із 41 великої агенції створили план Штучного інтелекту відповідно до вимог. Однак, і цей результат є дуже оптимістичний, оскільки, оскільки на практиці лише одне велике агентство, Департамент охорони здоров'я та соціальних служб (Department of Health and Human Services HHS), надав детальний план у відповідь [79].

30 жовтня 2023 р. президент Байден опублікував свій довгоочікуваний виконавчий Указ 14110 «Безпечна, захищена та надійна розробка та використання Штучного інтелекту» (Executive Order 14110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence). Це вважається одним із найвпливовіших заходів будь-якого уряду в світі щодо ШІ. Після випуску цього Указу США роблять значний крок в управлінні безпекою та ризиками ШІ. Потрібно відзначити, що Указ є дуже конкретним і набагато ширшим, ніж Указ 2020 р.— він містить близько 150 фактичних і конкретних вимог до агентств, назвавши 50 агентств, що свідчить про загальнодержавний підхід до регулювання ШІ [94].

Замість того, щоб далі впроваджувати ЕО 13859, адміністрація Байдена натомість переглянула тему регулювання Штучного інтелекту через проект Білля про права на Штучний інтелект (Blueprint for an AI Bill of Rights). Розроблений Управлінням науково-технічної політики Білого дому (Office of Science and Technology Policy | The White House), AIBoR включає детальний опис шкоди Штучного інтелекту економічним і громадянським правам, п'ять принципів пом'якшення цієї шкоди та пов'язаний перелік дій федеральних агентств.

AIBoR схвалює галузевий підхід до управління Штучним інтелектом із політичними втручаннями, адаптованими до окремих секторів, таких як охорона здоров'я, праця та освіта. Таким чином, його підхід значною мірою залежить від цих пов'язаних дій федерального агентства, а не від

централізованих дій, особливо тому, що AIBoR є необов'язковим керівництвом [109].

Те, що AIBoR безпосередньо не зобов'язує федеральні агентства пом'якшувати ризики Штучного інтелекту, стає зрозумілим з низки відповідей: одні агентства докладали значних зусиль і інші не реагували. Крім того, незважаючи на п'ять загальних принципів, викладених у AIBoR, більшість федеральних відомств можуть лише адаптувати свої вже існуючі правові повноваження до алгоритмічних систем. Найкраще це демонструють установи, що регулюють Штучний інтелект, який використовується для прийняття соціально-економічних рішень. Це включає Федеральну торгову комісію (FTC), яка може використовувати свої повноваження для захисту від «нечесних і оманливих» практик, щоб забезпечити правдивість реклами та деякі гарантії конфіденційності даних у системах ШІ.

Серед цих принципів адміністрація президента Байдена особливо активно виступала щодо расової рівності й у лютому 2023 р. опублікувала виконавчий наказ «Про подальше просування расової рівності та підтримку громад, які недостатньо обслуговуються через федеральний уряд» ("On Further Advancing Racial Equity and Supporting Underserved Communities Through the Federal Government" EO 14091). Другий виконавчий наказ щодо цього питання, EO 14091, наказує федеральним агентствам усунути нові ризики для громадянських прав, включаючи «алгоритмічну дискримінацію в автоматизованих технологіях». Поки рано знати про вплив цього нового розпорядження [92].

Якщо аналізувати Підхід ЄС до управління ризиками Штучного інтелекту характеризується, то потрібно відзначити, що він характеризується більш повним спектром законодавства, адаптованого до конкретних цифрових середовищ. ЄС планує ввести нові вимоги до Штучного інтелекту високого ризику в соціально-економічних процесах, використання Штучного інтелекту державою та регульованих споживчих товарів із системами Штучного

інтелекту. Інше законодавство ЄС забезпечує більшу публічну прозорість і вплив на розробку систем ШІ в соціальних медіа та електронній комерції.

Оскільки Європейський Союз має амбіцію стати світовим лідером у регулюванні Штучного інтелекту, в умовах швидкості технологічних змін і можливих викликів ЄС прагне до збалансованого підходу. В інтересах Союзу зберегти технологічне лідерство ЄС і забезпечити, щоб європейці могли отримати користь від нових технологій, розроблених і функціонуючих відповідно до цінностей, основних прав і принципів Союзу.

Для цього 19 лютого 2020 р. Єврокомісія опублікувала «Білу книгу про Штучний інтелект – європейський підхід до досконалості та довіри» (White Paper on Artificial Intelligence - A European approach to excellence and trust). Біла книга визначає варіанти політики щодо того, як досягти подвійної мети сприяння поширенню ШІ та усунення ризиків, пов'язаних із певним використанням такої технології. Ця пропозиція спрямована на реалізацію другої мети щодо розвитку екосистеми довіри, пропонуючи правову основу для надійного ШІ. Пропозиція ґрунтується на цінностях і фундаментальних правах ЄС і спрямована на те, щоб дати людям та іншим користувачам впевненість у прийнятті рішень на основі ШІ, водночас заохочуючи бізнес до їх розробки. ШІ має бути інструментом для людей і силою добра в суспільстві з кінцевою метою підвищення добробуту людей [58]. Після публікації Білої книги Комісія розпочала широкі консультації із зацікавленими сторонами, які значною мірою підтримали регуляторне втручання для вирішення проблем, викликаних дедалі більшим використанням ШІ.

Наступним кроком Єврокомісії стала розробка і прийняття нормативної бази щодо Штучного інтелекту з наступними конкретними цілями :

- забезпечити, щоб системи ШІ, що використовуються на ринку Союзу були безпечними та поважали чинне законодавство про основні права та цінності Союзу;

- забезпечити юридичну визначеність для сприяння інвестиціям та інноваціям у ШІ;
- покращити управління та ефективне забезпечення виконання чинного законодавства про основні права та вимоги безпеки, що застосовуються до систем ШІ;
- сприяти розвитку єдиного ринку для законних, безпечних і надійних програм ШІ та запобігати фрагментації ринку [59].

Парламент прийняв Закон «Про Штучний інтелект» у березні 2024 р., а Рада затвердила їх в травні 2024 р. Закон буде повністю застосовний через 24 місяці після набуття чинності, але деякі частини будуть застосовані раніше [55]. Основна мета — класифікувати різні системи ШІ відповідно до ризику, який вони несуть для користувачів. Різні рівні ризику означатимуть більшу чи меншу регуляцію. Після офіційного прийняття це буде перший у світі набір правил для ШІ. На практиці це перша всеосяжна постанова, яка враховує ризики Штучного інтелекту через набір зобов'язань і вимог, призначених для захисту здоров'я, безпеки та основних прав громадян ЄС і за його межами, і очікується, що вона матиме величезний вплив про глобальне управління ШІ [125].

Щоб відрізнити ШІ від більш простих програмних систем, ст. 3(1) Закону ЄС про Штучний інтелект визначає систему Штучного інтелекту як «машинну систему, яка розроблена для роботи з різними рівнями автономності та яка може демонструвати адаптивність після розгортання, і яка, для явних чи неявних цілей, робить висновки із вхідних даних які вона отримує, та може генерувати результати, такі як прогнози, вміст, рекомендації або рішення, які можуть впливати на фізичне або віртуальне середовище» [70].

Підхід ЄС до управління Штучного інтелекту є складним і багатогранним, він базується на чинному законодавстві, особливо на Загальному регламенті захисту даних (General Data Protection Regulation GDPR), і охоплює нещодавно прийняте законодавство, а саме Закон про

цифрові послуги (Digital Services Act (DSA) [93] та Закон про цифрові ринки (igital Markets Act (DMA)), а також законодавство, яке все ще активно діє: «ЄС свідомо розробив різні регуляторні підходи для різних цифрових середовищ, у кожному з яких різний ступінь акценту на ШІ» [99].

Окрім наслідків для конфіденційності даних, GDPR містить дві важливі статті, пов'язані з алгоритмічним прийняттям рішень. По-перше, GDPR стверджує, що алгоритмічним системам не можна дозволяти приймати важливі рішення, які впливають на законні права, без будь-якого контролю з боку людини. Згідно з цим пунктом, у 2021 р. Uber була зобов'язана відновити на роботі шістьох водіїв, які були звільнені виключно алгоритмічною системою компанії. По-друге, GDPR гарантує право особи на «значущу інформацію про логіку» алгоритмічних систем, що часом суперечливо вважається «правом на пояснення». На практиці такі компанії, як провайдери страхування житла, надають обмежені відповіді на запити щодо інформації щодо алгоритмічних рішень. Є багато відкритих питань щодо цього пункту, зокрема, як часто постраждалі особи запитують цю інформацію, наскільки ця інформація є для них цінною та що відбувається, коли компанії відмовляються надати інформацію [63].

Закон про цифрові послуги (DSA), прийнятий у листопаді 2022 р., розглядає Штучний інтелект як частину свого цілісного підходу до онлайн-платформ і пошукових систем. Створюючи нові вимоги до прозорості, вимагаючи незалежних аудитів і дозволяючи незалежні дослідження на великих платформах, DSA розкриє багато нової інформації про функції та шкоду ШІ на цих платформах. Крім того, Закон вимагає, щоб великі платформи пояснювали свій Штучний інтелект щодо рекомендацій вмісту, наприклад заповнення стрічок новин, і пропонували користувачам альтернативну систему рекомендацій, яка не базується на конфіденційних даних користувачів. Якщо ці системи рекомендацій сприяють поширенню дезінформації, а великі платформи

не можуть пом'якшити цю шкоду, їм можуть загрожувати штрафи відповідно до DSA [51].

Незважаючи на те, що його часто називають «горизонтальним», Закон про Штучний інтелект реалізує багаторівневу систему нормативних зобов'язань для конкретного переліку програм ШІ. Закону встановлює зобов'язання для провайдерів і користувачів відповідно до рівня ризику, який представляє ШІ. Ці рівні поділяються на чотири різні категорії (ризик): I. Неприйнятний ризик – системи ШІ з неприйнятним рівнем ризику для «безпеки людей».

Системи такого рівня ризику :

- Когнітивне поведінкове маніпулювання людьми або окремими вразливими групами;
- Соціальна оцінка: класифікація людей на основі поведінки, соціально-економічного статусу чи особистих характеристик;
- Системи біометричної ідентифікації в реальному часі та віддалені, наприклад розпізнавання обличчя. Правове регулювання таких систем заборонено [91].

Участь у заборонених практиках ШІ може призвести до штрафів у розмірі до 40 мільйонів євро (остаточна сума 35 млн.) або до 7% глобального річного обороту компанії, залежно від того, що більше. Однак штрафи будуть «пропорційними» та враховуватимуть ринкову позицію невеликих провайдерів, припускаючи, що можуть існувати більш гнучкі правила для стартапів [125]. Використання Штучного інтелекту з високим ризиком включають критичну інфраструктуру, освіту та професійну підготовку, працевлаштування, основні приватні та державні послуги (наприклад, охорона здоров'я, банки), певні системи правоохоронних органів, управління міграцією та кордонами, правосуддя та демократичні процеси (наприклад, вплив на вибори) . Такі системи повинні оцінювати та зменшувати ризики, вести журнали використання, бути прозорими та точними та забезпечувати людський контроль стартапів [101].

II група високого ризику - системи, які негативно впливають на безпеку або основні права. Системи такого рівня ризику :

Системи Штучного інтелекту, які використовуються в продуктах, які підпадають під дію законодавства ЄС щодо безпеки продуктів. Це включає в себе іграшки, авіацію, автомобілі, медичне обладнання та ліфти. Системи Штучного інтелекту, що належать до певних сфер, які повинні бути зареєстровані в базі даних ЄС:

- Управління та експлуатація критичної інфраструктури
- Освіта та професійна підготовка
- Працевлаштування, управління працівниками та доступ до самозайнятості
- Доступ і користування основними приватними та державними послугами та перевагами
- Правоохоронні органи
- Управління міграцією, притулком та прикордонним контролем.

Допомога в правовому тлумаченні та застосуванні законодавства. Правове регулювання таких систем дозволено після оцінки до та після виходу на ринок.

Постачальники Штучного інтелекту високого ризику повинні будуть зареєструвати свій Штучний інтелект у базі даних ЄС, якою керує Комісія, перш ніж вивести його на ринок. Постачальникам з-за меж ЄС потрібен буде уповноважений представник у ЄС для демонстрації відповідності та постмаркетингового моніторингу.

3. Обмежений ризик. Системи, що становлять такий рівень ризику : системи Штучного інтелекту, які генерують або маніпулюють зображеннями, аудіо-, відеовмістом і чат-ботами.

Правове регулювання таких систем : дозволено з мінімальними вимогами прозорості. Після взаємодії з додатками користувач може вирішити, чи хоче він продовжувати ним користуватися. Користувачі також повинні знати, коли вони

взаємодіють з ШІ. Ці системи Штучного інтелекту повинні відповідати вимогам прозорості, таким як розкриття інформації про те, що контент створено ШІ; розробка моделі для запобігання створенню нею незаконного вмісту та публікація резюме захищених авторським правом даних, які використовуються для навчання.

4. Мінімальний (низький) ризик. Системи, що становлять такий рівень ризику : генеративний Штучний інтелект (наприклад, ChatGPT)

Правове регулювання таких систем: дозволено без зобов'язань, але, ймовірно, є добровільними кодексами поведінки [91].

Покарання за порушення деяких інших положень підлягають максимальному штрафу в розмірі 15 мільйонів євро або 3 відсотки світового річного обороту, залежно від того, яка сума є більшою. Максимальний штраф за надання невірної, неповної або оманливої інформації нотифікованим органам або національним компетентним органам становить 7,5 мільйонів євро або 1 відсоток річного обороту в усьому світі, залежно від того, яка сума є більшою (ст. 99(5) Закону ЄС про AI). Для малих і середніх підприємств та новостворених підприємств штрафи за все вищезазначене підлягають однаковим максимальним відсоткам або сумах, але залежно від того, яка з них нижча (ст. 99(6) Закону ЄС про AI) [91].

Десятки провідних лідерів європейського бізнесу відмовилися від підтримки запропонованого ЄС закону про Штучний інтелект, попередивши, що він може зашкодити конкурентоспроможності ЄС і призвести до відтоку інвестицій. Вони стверджують, що проект правил заходить занадто далеко, зокрема в регулюванні генеративних моделей ШІ, технології, що лежить в основі популярних платформ, таких як ChatGPT.

Досить цікавим є нормативна база Великої Британії в регулюванні ШІ.

Прем'єр-міністр Великої Британії Ріші Сунак та його уряд зайняли агресивну позицію щодо регулювання Штучного інтелекту, прагнучи очолити

світові регулятори, надаючи сприятливу для бізнесу відповідь на скептицизм ЄС.

Як повідомляється, 3 червня Сунак висловив президенту Байдену бажання створити глобальний орган зі спостереження за Штучним інтелектом за зразком Міжнародного агентства з атомної енергії, але зі штаб-квартирою в Лондоні. Нещодавно він також написав у Твіттері: «Зроблений безпечно та безпечно, Штучний інтелект має потенціал для трансформації та розвитку економіки». У квітні Сунак оголосив про створення цільової групи зі 100 мільйонами фунтів стерлінгів для «розробки безпечного та надійного використання ... ШІ ... в економіці». Це оголошення відбулося після публікації британського уряду від 29 березня під назвою «Проінноваційний підхід до регулювання ШІ» [26].

Велика Британія зазначила, що «уникатиме жорсткого законодавства» у цій сфері та надаватиме повноваження існуючим регуляторам замість створення нового окремого регулятора. Він також «спочатку ... [не] має наміру запроваджувати нове законодавство» [26], побоюючись перевантажити бізнес. Біла книга була відкрита для коментарів до 21 червня, і уряд Великобританії аналізує надані відгуки з наміром випустити «Дорожню карту регулювання ШІ». Тим часом Сунак виступив із низкою пропозицій, які він планує обговорити з президентом Байденом. Пропозиції Сунака щодо глобального лідерства включають згадану вище пропозицію щодо глобального органу Штучного інтелекту під керівництвом Лондона, іншу пропозицію щодо міжнародного дослідницького органу Штучного інтелекту в Лондоні та міжнародний саміт Штучного інтелекту в Лондоні цієї осені. Хоча біла книга є ключем до розуміння поточної позиції Великої Британії, глобальна боротьба за владу вводить ключові динаміки, які можуть суттєво сформувати наступні кроки Великої Британії.

Біла книга дотримується рекомендації Департаменту цифрових технологій, культури, медіа та спорту від липня 2022 року, яка вказує на те, що

статичного визначення може бути недостатньо для охоплення поточних і майбутніх застосувань ШІ. Чи буде це визначення більш чи менш широким, ніж визначення Закону ЄС про АІ, ймовірно, залежатиме від його застосування на практиці [26].

Що стосується оцінки ризиків, Велика Британія вирішила застосувати підхід до регулювання «залежно від контексту», розглядаючи «результати, які Штучний інтелект, ймовірно, створить у конкретних програмах». Іншими словами, ризики не будуть віднесені до конкретних технологій чи секторів. Натомість регулятори розглядатимуть загальні результати та зважуватимуть їх проти альтернативних витрат.

Країна, відома своїм повсюдним використанням відеоспостереження, здається, готова прийняти нові виклики ШІ, але довіра до існуючих регуляторів і процесів викликає практичні питання: «Як Офіс уповноваженого з питань інформації, Управління фінансового контролю та Комісія з питань рівності та прав людини співпрацюватимуть з часом? Чи матимуть вони ресурси та досвід для роботи з технологією, що розвивається? До честі біла книга містить пропозицію щодо централізованих функцій, у тому числі центральної функції ризиків, з повноваженнями, які, здається, включають координаційні та ліквідаційні ролі, тоді як регулятори будуть «визначати та визначати пріоритети нових ризиків ШІ у своєму секторі» [26].

Але ризики, пов'язані з цією технологією, не поширюються чітко на секторні лінії. Однак несправедливо припускати, що уряд не врахував цього. У білій книзі чітко зазначено: «Створення нового міжгалузевого регулятора, що спеціалізується на Штучному інтелекті, створить складність і плутанину, що підриває та, ймовірно, суперечить роботі наших існуючих експертних регуляторів» [26].

Ще однією проблемою для здатності Великої Британії лідирувати в регулюванні Штучного інтелекту є її виключення із відносин США та ЄС після

Brexit і, як наслідок, передові дискусії про співпрацю між двома сторонами [48].

В Україні вже розпочато створення правової бази для регулювання Штучного інтелекту. У 2020 році була прийнята Концепція розвитку Штучного інтелекту [13], яка вперше на законодавчому рівні визначила мету, принципи та завдання розвитку цієї технології в країні.

Концепція є прогресом для формування інноваційної економіки України. Вона закладає основу для міжнародного співробітництва у сфері ШІ та гармонізації законодавства з ЄС. Протест успіху її реалізації залежить від співпраці уряду, наукової спільноти, бізнесу та громадянського суспільства, а також від відповідного фінансування та координації.

Цей документ передбачає програму України створити сприятливі умови для розвитку технологій ШІ, сприяючи одночасно захисту прав людини та підтриманню економічного розвитку.

Додатково, Національна асоціація юристів України утворила Робочу групу з правового регулювання Штучного інтелекту. Ця група займається аналізом ключових правових аспектів розробки і використання Штучного інтелекту, визначенням меж його застосування у різних сферах, а також захистом персональних даних, що збираються системами ШІ, включаючи правила їх збереження та використання. Робоча група також працює над розвитком галузі та забезпеченням захисту прав людини у контексті використання ШІ [35].

У грудні 2022 р. в Україні були внесені зміни до закону про авторське право, які стосуються прав на твори, створені Штучним інтелектом. Закон тепер визнає особливі права «*sui generis*» на неоригінальні об'єкти, створені без участі людини.

Твори, створені Штучним інтелектом, які відрізняються від подібних творів і створені без участі людей, підлягають цим правам *sui generis*. Права включають можливість використовувати твір і дозволяти або забороняти

використання третіми особами. Матеріальні права *sui generis* належать фізичній особі, яка має ліцензію на програмне забезпечення, яка створила твір, або власнику програмного забезпечення, як зазначено в окремих угодах.

Термін дії прав *sui generis* на твори, створені ШІ, становить 25 років з першого дня року, наступного за роком створення. Крім того, неоригінальні бази даних тепер користуються правами *sui generis* протягом 15 років з моменту їх створення [81, с. 192].

Для регулювання Штучного інтелекту в Україні Міністерство цифрової трансформації підготувало Білу книгу 2024 року. У цьому документі детально викладено підходи до регулювання ШІ в Україні, що дозволить компаніям підготуватися до майбутніх законодавчих змін і забезпечити безпеку своїх продуктів для громадян.

Розроблений підхід враховує інтереси всіх ключових сторін — громадян, компаній і держави. Основна мета полягає в захисті прав українців, водночас підтримуючи розвиток інновацій. Це дозволить українським компаніям підвищити свою конкурентоспроможність і вийти на міжнародні ринки. Для держави ж це можливість інтегруватися в ЄС, узгоджуючи законодавство у сфері Штучного інтелекту з європейськими стандартами.

«Жорсткі заходи щодо інновацій можуть загальмувати промисловий розвиток. Тому ми обрали м'який підхід до створення регулювання в Україні, що ґрунтується на принципах «знизу вгору». Спочатку ми підготуємо те, що потрібно компаніям, а потім приймемо законодавство, аналогічне європейському закону про Штучний інтелект. На першому підготовчому етапі компанії заздалегідь отримуватимуть від держави інструменти для підготовки до майбутнього законодавства» [20] - каже заступник міністра цифрової трансформації Олександр Борняков.

Інструменти, які держава надасть компаніям, міститимуть загальні та галузеві рекомендації з різних напрямків і аспектів використання ШІ, починаючи від освіти і журналістики і закінчуючи порадами щодо обробки

персональних даних. Крім рекомендацій, буде створено добровільний кодекс поведінки, платформу правової підтримки компаній і регуляторну пісочницю для перевірки відповідності високотехнологічних продуктів майбутнім вимогам.

Білу книгу було підготовлено за підтримки USAID «Кібербезпека для критичної інфраструктури в Україні» та «Цифрової трансформації» за підтримки USAID та UK Dev [20].

У відповідь на сучасні виклики, проект EU4DigitalUA, фінансований ЄС, разом із партнерами з Офісу омбудсмена та Міністерства цифрової трансформації України, представив перші в Україні рекомендації щодо захисту прав людини та приватності у сфері Штучного інтелекту (ШІ) [56].

Документ під назвою «Права людини в епоху Штучного інтелекту: виклики та правове регулювання», представлений у лютому 2024 року, аналізує вплив ШІ та обговорює питання управління цими технологіями відповідно до національних і міжнародних стандартів. Він також надає загальні рекомендації щодо вирішення можливих проблем і використання потенційних можливостей ШІ. Як зазначила Юлія Деркаченко, представниця Уповноваженого з прав інформації Секретаріату Уповноваженого з прав людини: «Розуміння принципів, яких дотримуються розробники та компанії, що працюють з ШІ, є важливим для формування довіри до електронних сервісів в Україні. Рекомендації сприятимуть підвищенню обізнаності та запобіганню порушень прав на конфіденційність» [56].

Ці рекомендації є важливим кроком до комплексного регулювання ШІ в Україні. Наступні етапи включатимуть розробку більш конкретних рекомендацій і методологій для оцінки впливу ШІ на права людини, створення добровільних кодексів поведінки для бізнесу, регуляторних пісочниць для тестування ідей, а також публікацію Білої книги з регулювання Штучного інтелекту для забезпечення належного захисту прав людини у контексті ШІ [56].

Отже, попри прогрес, регулювання ШІ залишається проблемою через швидкий розвиток технологій і складність їх правового регулювання. Україна робить акцент на міжсекторальному співробітництві, залучаючи уряд, бізнес, науковців і громадськість. Такий підхід створює сприятливі умови для інновацій, не обмежуючи розвиток промисловості жорсткими нормативними бар'єрами.

Етика, прозорість і безпека — ключові принципи, які інтегруються у вітчизняну стратегію. Водночас зростає потреба у більш систематичному підході, який викликає виклики кібербезпеки, алгоритмічної дискримінації та захисту конфіденційності.

Узгодження національних норм із міжнародними стандартами сприятиме інтеграції України в глобальні ринки, підтримуючи цю демократичну цінність та верховенство права. Сучасний фокус на співпраці з міжнародними партнерами, включаючи USAID та ЄС, є запорукою ефективного управління ШІ в майбутньому.

3.2. Інструменти та механізми міжнародного -правового регулювання систем Штучного інтелекту

Міжнародне організаційно-правове регулювання системи штучного інтелекту (ШІ) стає пріоритетним для світового співтовариства, оскільки технології ШІ мають транскордонний характер і впливають на глобальні виклики. Міжнародні організації, такі як ООН, Рада Європи та Організація економічного співробітництва та розвитку (ОЕСР), формують рамкові політики для забезпечення відповідного використання ШІ. Ці інструменти включають розробку стандартів, рекомендацій, а також укладення міжнародних угод, спрямованих на захист прав людини, етичність алгоритмів і технологічну безпеку.

Серед основних механізмів міжнародного регулювання — створення кодексів етики, технічних стандартів, моніторингових систем і платформи для співпраці між державами. Особливу увагу приділяють гармонізації національних підходів до регулювання, щоб уникнути правових прогалин, які можуть використовуватися для недобросовісного застосування ШІ. Такий підхід сприяє зміцненню довіри між країнами, бізнесом і суспільством, забезпечуючи баланс між інноваціями та захистом загальних інтересів.

Березень 2024 р. став вершиною свого роду змагань щодо того, яка міжнародна організація першою прийме інструмент, що намагається регулювати розробку, виробництво та використання Штучного інтелекту (ШІ).

11 березня Генеральна Асамблея ООН одноголосно прийняла резолюцію під назвою «Використання можливостей безпечних, захищених і надійних систем Штучного інтелекту для сталого розвитку» [19]. Відповідно до відповідної веб-сторінки ООН, це «визначна резолюція», представлена таким коментарем: «Прийнявши проект резолюції, очолюваний Сполученими Штатами, без голосування, Асамблея також підкреслила повагу, захист і підтримку прав людини в дизайні, розробка, розгортання та використання ШІ. Текст був підтриманий понад 120 іншими державами-членами. Генеральна Асамблея також визнала потенціал систем Штучного інтелекту для прискорення та забезпечення прогресу на шляху до досягнення 17 цілей сталого розвитку. Це перший випадок, коли Асамблея прийняла резолюцію щодо регулювання нової сфери. Повідомляється, що раніше цього місяця радник США з національної безпеки заявив, що прийняття стане «історичним кроком вперед» для безпечного використання ШІ» [63].

Відповідно до останнього пункту (№ 13) Резолюції, Генеральна Асамблея: «Визнає, що система Організації Об'єднаних Націй, відповідно до свого мандату, унікально сприяє досягненню глобального консенсусу щодо безпечних, захищених і надійних систем Штучного інтелекту, що узгоджується з міжнародне право, зокрема Статут ООН, Загальна декларація прав людини і

Порядок денний сталого розвитку на період до 2030 року, в тому числі шляхом сприяння інклюзивному міжнародному співробітництву та сприянню залученню, участі та представництву країн, що розвиваються, у обговореннях» [19]; у пункті № 12 тексту сказано, що він «також з нетерпінням чекає на загальний огляд Генеральною Асамблеєю у 2025 році прогресу, досягнутого після Всесвітнього саміту з питань інформаційного суспільства» [19].

Шукаючи більше змісту, ніж резолюція ООН у середині березня, краще взяти до уваги позитивне голосування Європейського парламенту – 523 голоси «за», 46 «проти» та 49 «утрималися» – щодо запропонованого Регламенту Європейського Союзу щодо Європейського парламенту та Ради про встановлення узгоджених правил Штучного інтелекту (Закон про Штучний інтелект) та внесення змін до певних законодавчих актів Союзу, щодо яких було досягнуто політичну угоду під час сесії «трилогу» між представниками Європейського парламенту та Ради ЄС [34]. Як слушно нагадується на відповідному Інтернет-сайті Європейського парламенту [32], текст «все ще підлягає остаточній юрист-лінгвістичній перевірці та, як очікується, буде остаточно ухвалений до закінчення легіслатури (через так звану процедуру виправлення)» [32] та «також має бути офіційно схвалено Радою» [32]. Остаточне формулювання тексту буде доступним лише після його публікації в офіційному журналі ЄС (протягом 2024 року), проект AI Act вже був предметом великої кількості літератури, в тому числі в цьому журналі [52].

Поки інституції Європейського Союзу працювали над регулюванням ШІ, над цим питанням також працювала Рада Європи, яка об'єднує всі європейські держави, крім Білорусі та Росії [16]. Рада Європи (РЄ) була заснована Лондонським договором 5 травня 1949 року. Цей договір підписали 10 європейських держав і він набув чинності 3 серпня 1949 року. Це найстаріша з організацій, створених після Другої світової війни з метою об'єднати європейські країни, які поділяють цінності ліберальної демократії.

14 березня 2024 року Комітет Ради Європи зі Штучного інтелекту (CAI) прийняв проект Рамкової конвенції про Штучний інтелект, права людини, демократію та верховенство права [44] (далі FCIA), який буде подано до Комітету міністрів, який, швидше за все, затвердить його, таким чином відкриваючи процедуру для підписання або приєднання держав-членів Ради Європи і третіх держав [25].

Було б неправильно сказати, що дві організації працювали паралельно: як свідчить сам текст, існувала і продовжує існувати велика взаємодія між інституціями двох європейських організацій [12]. Це найменше, що можна було зробити, враховуючи, що всі 27 держав-членів ЄС також є членами Ради Європи разом із 19 іншими європейськими державами.

Згідно з проектом пояснювальної доповіді FCIA [18], у грудні 2021 року Комітет міністрів Ради Європи ухвалив рішення «дозволити участь Європейського Союзу та зацікавлених неєвропейських країн, які поділяють цінності та мету Ради Європи – таких як Аргентина, Австралія, Канада, Коста-Ріка, Святий Престол, Ізраїль, Японія, Мексика, Перу, Сполучені Штати Америки та Уругвай – у переговорах по CAI та участь у розробці цієї Рамкової конвенції як держав-спостерігачів» [18].

Саме тому 21 листопада 2022 року Рада ЄС ухвалила рішення про початок переговорів з ЄС щодо Конвенції Ради Європи про Штучний інтелект, права людини, демократію та верховенство права [21]. Деякі тексти рішення варто процитувати.

«Союз прийняв загальні правила, на які впливатимуть елементи, що розглядаються стосовно конвенції. Ці елементи включають, зокрема, комплексний набір правил у сфері єдиного ринку для продуктів і послуг, для яких можна використовувати системи Штучного інтелекту, а також вторинне законодавство Союзу, яке імплементує Хартію ЄС про основні права, враховуючи, що ці права, ймовірно, будуть піддаватися негативному впливу за певних обставин розробкою та використанням певних систем ШІ» [34].

Декларація зазначає, що сфера застосування, передбачена Конвенцією, і сфера застосування пропозиції АІА «значною мірою збігаються з цією законодавчою пропозицією за сферою застосування, оскільки обидва інструменти мають на меті встановити правила, застосовні до проектування, розробки та застосування систем Штучного інтелекту, які надаються та використовуються державними або приватними організаціями» [32]. Крім того, відповідно до пункту 6: «отже, укладення конвенції може вплинути на існуючі та передбачувані майбутні спільні правила Союзу або змінити сферу їх застосування у значенні статті 3 (2) Договору про функціонування Європейського Союзу (ДФЕС)» [32]. Досить вражаючим є те, що цей текст посилається на статтю 3(2) про цінності Союзу, згідно з якою «Союз запропонує своїм громадянам простір свободи, безпеки та справедливості без внутрішніх кордонів, у якому забезпечується вільне пересування людей у поєднанні з відповідними заходами щодо контролю зовнішніх кордонів, притулку, імміграції та запобігання та боротьби зі злочинністю» [32]. Як ми побачимо, пропозиція Комісії від 21 квітня 2021 року [95], а також текст, схвалений Європейським парламентом у березні 2024 року [73], стосуються лише правових основ, що стосуються внутрішнього ринку, а не тих, що стосуються контроль на внутрішніх і зовнішніх кордонах Союзу.

Слід також зазначити, що Європейський інспектор із захисту даних (EDPS) опублікував звіт щодо проекту FCIA, на який посилається у примітці до Рішення Ради, з рядом рекомендацій, які часто повторювалися в наступних версіях проекту [4]. У своїх загальних коментарях ЄКЗД зазначає, що «ринково-орієнтований підхід узгоджується з однією з головних цілей запропонованого Закону про Штучний інтелект – єдиний ринковий вимір регулювання систем ШІ. У той же час повноваження Ради Європи є набагато ширшими. На цьому фоні ЄКЗД вважає, що конвенція являє собою важливу можливість доповнити запропонований Закон про АІ шляхом посилення захисту основних прав усіх осіб, які постраждали від системи ШІ. Відповідно та згідно зі Спільним

висновком щодо Закону про Штучний інтелект, ЄКЗД вважає, що захист прав окремих осіб і груп осіб, які підлягають використанню систем Штучного інтелекту, має надаватися більшого значення серед загальних цілей для переговорів щодо конвенції» [74]. Як ми побачимо, FCIA включила спеціальні положення для Європейського Союзу з версії від грудня 2023 року, які, очевидно, є результатом участі Комісії в переговорах.

AIA встановлює основи та структури безпечної та надійної екосистеми Штучного інтелекту, а Конвенція зосереджується на його впливі на людей і демократичне суспільство. AIA є методичним, детальним і точним, прокладаючи чіткий шлях через технічну та юридичну складність, встановлюючи тверді стандарти та конкретні зобов'язання для постачальників і користувачів або розробників систем ШІ. На противагу цьому, з ліричного боку, Конвенція піднімається до нормативної інтеграції фундаментальних цінностей, етичних принципів і прав людини, якими має керуватися еволюція ШІ» [78].

Розділ II FCIA присвячений «Загальним зобов'язанням». Стаття 4 зазначає, що «Кожна Сторона повинна вжити або підтримувати необхідні заходи для забезпечення відповідності діяльності, пов'язаної з життєвим циклом систем Штучного інтелекту, її зобов'язанням щодо захисту прав людини відповідно до міжнародного та національного права» [19]. Це передбачає не лише прийняття необхідних нормативно-правових актів і законодавства (частково охопленого AIA для держав-членів ЄС), але й забезпечення необхідних людських та бюджетних ресурсів, а також проведення навчання та інформаційних кампаній. Як зазначено в проекті пояснювальної доповіді: «Це зобов'язання результату, а не зобов'язання засобів. В цьому контексті принцип субсидіарності має ключове значення, оскільки несе відповідальність за забезпечення дотримання прав людини та відшкодування за порушення» [103].

Стаття 5 вказує, що «заходи повинні бути спрямовані на запобігання використанню систем Штучного інтелекту для підриву цілісності, незалежності та ефективності демократичних інститутів і процесів, зокрема принципу поділу влади, незалежності судової влади та доступу до правосуддя» [19]. Також зазначено, що «Кожна Сторона повинна вжити або підтримати заходи, спрямовані на захист своїх демократичних процесів у межах життєвого циклу систем Штучного інтелекту, забезпечуючи справедливий доступ осіб до публічних дебатів та участь у них, а також можливість вільно формувати свої погляди» [19].

FCIA встановлює зобов'язання вживати заходів щодо «цілісності демократичних процесів і поваги до верховенства права» (стаття 5), а також «людської гідності та індивідуальної автономії» (стаття 7). Як ми побачимо нижче, FCIA є частиною основної місії Ради Європи, яка полягає в захисті прав людини та верховенства права в демократичному суспільстві, насамперед за допомогою обов'язкових правових інструментів, як це передбачено Статутом Ради Європи та ЄСПЛ.

Глава III присвячена «Принципам, пов'язаним з діяльністю в рамках життєвого циклу систем Штучного інтелекту», в яких «встановлюються загальні принципи, які кожна Сторона має впроваджувати щодо систем Штучного інтелекту відповідно до її внутрішньої правової системи та іншої зобов'язання цієї Конвенції» (стаття 6). Це «Людська гідність та індивідуальна автономія» (стаття 7); «Прозорість та нагляд» (стаття 8); «Звітність та відповідальність» (ст. 9); «Рівність і недискримінація» (стаття 10); «Захист конфіденційності та особистих даних» (стаття 11) «Надійність», тобто «заходи щодо підвищення надійності систем Штучного інтелекту та довіри до їх результатів, які можуть включати вимоги, пов'язані з належною якістю та безпекою протягом життєвого циклу систем Штучного інтелекту»; та «Безпечні інновації», тобто «покликані забезпечити, у відповідних випадках, створення контрольованого середовища для розробки, експериментування та тестування

систем Штучного інтелекту під наглядом своїх компетентних органів» (стаття 13).

Доктор Лоренцо Котіно Уесо, професор конституційного права в Університеті Валенсії каже, що «Конвенція має не лише символічну та мета-юридичну цінність, але також є нормативним інструментом, здатним до квазі-конституційної інтеграції в правові системи держав-учасниць і має великий інтерпретаційний потенціал. Ось чому Конвенція ІА перевершує десятки декларативних і м'яких правових інструментів, які і так були зайвими, нешкідливими і навіть нудними» [80].

Глава IV присвячена «засобам правового захисту» та «процесуальним гарантіям». Це зобов'язання держав-учасниць, а не система засобів правового захисту на рівні РЄ, як ми побачимо нижче. У розділі V йдеться про «Оцінку та пом'якшення ризиків і несприятливих впливів».

Глава VI присвячена «Імплементатії Конвенції» з регулярними положеннями в останніх інструментах Ради Європи щодо недискримінації (стаття 17), прав людей з обмеженими можливостями та дітей (стаття 18), громадських консультацій (стаття 19), гарантій для існуючі права людини (стаття 21) і більш широкий захист (стаття 22). Стаття 20 «Цифрова грамотність і навички» більш конкретно стосується Штучного інтелекту: «Кожна Сторона повинна заохочувати та сприяти належній цифровій грамотності та цифровим навичкам для всіх верств населення, включаючи спеціальні експертні навички для осіб, відповідальних за ідентифікацію, оцінку, запобігання та пом'якшення наслідків. ризиків, створених системами Штучного інтелекту» [19].

Глава VII присвячена «Механізму подальших заходів і співпраці». Що стосується розділу VIII щодо заключних положень, важливо, що лише п'ять держав мають ратифікувати, з яких принаймні три мають бути членами Ради Європи, що свідчить про намір активувати Конвенцію ІА якомога швидше.

Доктор Лоренцо Котіно Уесо зазначає: «Хоча загалом Конвенція ІА не характеризується чіткими конкретними зобов'язаннями та правами, є кілька причин, щоб прийняти її до уваги нормативно. Я вважаю доречним положення Конвенції ІА про загальні «принципи», застосовні до всіх систем ІА. У зв'язку з цим варто нагадати, що протягом багатьох років серед десятків декларацій і документів деякі основні етичні принципи Штучного інтелекту були помітні та виділені. Гарвард проаналізував понад тридцять основних міжнародних і корпоративних декларацій щодо етики Штучного інтелекту та об'єднав їх у конфіденційність, підзвітність, безпеку, прозорість і пояснюваність, справедливість і недискримінацію, людський контроль, професійну відповідальність, людські цінності та сталість. Майбутня Конвенція є позитивною в тому, що вона виходить за рамки декларацій у сфері м'якого права та регулює ці принципи, якщо можна сказати, вона переходить від музики етики до театру права. Проте є деякі конкретні елементи Конвенції, які можуть дещо виходити за рамки ЗІС та законодавства ЄС» [77].

Нарешті, необхідно представити конкретні положення, що виникли в результаті участі Комісії ЄС у переговорах, щодо особливостей підписання та ратифікації договорів Ради Європи на відміну від ухвалення регламенту чи директиви ЄС.

По-перше, відповідно до статті 27 – Наслідки Конвенції:

«1. Якщо дві або більше Сторін уже уклали угоду чи договір щодо питань, які охоплюються цією Конвенцією, або іншим чином врегулювали свої відносини з цих питань, вони мають право продовжувати застосовувати цю угоду чи договір або регулювати свої відносини відповідно до них, за умови, що це не суперечить меті та завданням цієї Конвенції.

2. Сторони, які є членами Європейського Союзу, у своїх відносинах застосовують норми Європейського Союзу, що регулюють питання, які входять до сфери дії цієї Конвенції, при цьому не порушуючи мету та завдання Конвенції та не заважаючи її повному застосуванню з інших сторін. Це також

стосується інших Сторін у межах, в яких вони підпорядковані відповідним правилам [45].

Версія, запропонована до останнього засідання САІ у березні 2024 року, також містила спеціальне положення для ЄС у статті 29 – Врегулювання суперечок: «У разі суперечки між Сторонами щодо тлумачення або застосування цієї Конвенції, яка не може бути вирішена шляхом Конференція Сторін, як це передбачено в пункті 1 е статті 24, вони прагнуть вирішити спір шляхом переговорів або будь-яким іншим мирним способом на свій вибір. Європейський Союз та його держави-члени у своїх відносинах одна з одною не можуть використовувати статтю 29 Конвенції. Держави-члени Європейського Союзу також не можуть використовувати цю статтю Конвенції, оскільки спір між ними стосується тлумачення або застосування права Європейського Союзу» [45]. Останні два речення не увійшли до редакції, прийнятої 14 березня фактично були нагадуванням про загальновідомі принципи права ЄС.

Спеціальний комітет зі Штучного інтелекту (САНАІ) РЄ, створений 11 вересня 2019 року, було перейменовано в Комітет зі Штучного інтелекту (САІ), в січні 2022 року.

Мандат САІ [100] було прийнято на період з 1 січня 2022 року до 31 грудня 2024 року в рамках програми «Ефективна імплементація ЄКПЛ», що пояснює фокус FCIA на верховенстві права та правах людини. Мандат було прийнято під керівництвом Комітету міністрів (CoM), у якому держави-члени Ради Європи, як правило, представлені їхніми постійними представниками в Страсбурзі, у виняткових випадках – їхніми міністрами закордонних справ. Рада Міністрів може ухвалювати резолюції та, зокрема, рекомендації для урядів держав-членів, зокрема щодо подальших заходів, які мають бути вжиті у зв'язку з рішеннями Європейського суду з прав людини, які є обов'язковими для держав (ЄКПЛ, стаття 46). Статут РЄ визначає процедуру голосування (стаття 20). Вони коливаються від більшості представників з одностайністю голосів, поданих для найважливіших питань, до більшості представників з

більшістю у дві третини голосів, поданих для більшості резолюцій, до простої більшості для питань, що стосуються Регламенту або фінансові та адміністративні правила.

Комітет доручив CAI [100] «взяти до уваги відповідні ключові висновки та виклики, викладені у звіті Генерального секретаря про стан демократії, прав людини та верховенства права за 2023 рік «Запрошення підтвердити зобов'язання до цінностей і стандартів Рада Європи». Метою було «започаткувати міжнародний переговорний процес і провести роботу для завершення відповідної правової бази щодо розробки, проектування, використання та виведення з експлуатації Штучного інтелекту на основі стандартів Ради Європи щодо прав людини, демократії та верховенства права та інші відповідні міжнародні стандарти та сприяють інноваціям, які можуть складатися із обов'язкового правового інструменту поперечного характеру, включаючи, зокрема, загальні принципи, а також додаткові зобов'язуючі чи необов'язкові інструменти для вирішення проблем, пов'язаних із застосуванням Штучного інтелекту у конкретних секторах, згідно з відповідними рішеннями Комітету міністрів» [100]. Він також мав «зберігати наскрізний підхід, також координуючи свою роботу з іншими міжурядовими комітетами та організаціями Ради Європи, які однаково розглядають наслідки Штучного інтелекту в їхній відповідній сфері діяльності, надаючи цим комітетам та організаціям керівництво відповідно до з розробкою законодавчої бази та надання їм допомоги у вирішенні проблем» [100], а також «базувати роботу на переконливих доказах і інклюзивному процесі консультацій, у тому числі з міжнародними та наднаціональними партнерами, щоб забезпечити глобальний погляд на предмет» [100]. Нарешті, метою було «внести свій внесок у досягнення Порядку денного сталого розвитку ООН на період до 2030 року та переглянути прогрес у його досягненні, зокрема щодо Цілі 5 - Гендерна рівність, Цілі 16 - Мир, справедливість і міцні інституції» [100].

Як резюмував Доктор Лоренцо Котіно Уесо: «цей мандат мав чіткий намір подолати кордони, прагнучи створити «інструмент, привабливий не лише для держав Європи, але й для якомога більшої кількості держав з усіх регіонів світу» [77], залучаючи «спостерігачів», таких як Ізраїль, Канада, Сполучені Штати, Японія, Глобальне партнерство зі Штучного інтелекту (GPAI), Інтернет-компанії та організації громадянського суспільства» [74].

Відповідно до статті 30 (1) FCIA, «1. Ця Конвенція відкрита для підписання державами-членами Ради Європи, державами, які не є членами Ради Європи, але брали участь у її розробці, а також Європейським Союзом». Європейський Союз є учасником кількох договорів Ради Європи, зокрема Стамбульської конвенції про запобігання насильству щодо жінок і домашньому насильству, яка набула чинності 1 січня 2023 року, а його приєднання до ЄКПЛ регулюється статтею 17 Протоколу 14 до ЄКПЛ, що змінює систему моніторингу Конвенції, а також статтею 16 ДФЄС [90].

Згідно зі статтею 31 (1) FCIA «1. Після набрання чинності цією Конвенцією Комітет міністрів Ради Європи може, після консультацій із Сторонами та отримання їхньої одностайної згоди, запросити будь-які держави, що не є членами Ради Європи і не брали участі в розробці Конвенції, приєднатися до неї. Таке рішення має бути прийняте більшістю голосів, зазначеною у статті 20.d Статуту Ради Європи, а також одностайним голосуванням представників Сторін, що мають право засідати в Комітеті міністрів». Декілька неєвропейських держав приєдналися до договорів РЄ, зокрема Канада, Чилі, Коста-Ріка, Святий Престол, Японія, Мексика та США. Щодо Конвенції Ради Європи про відмивання, пошук, арешт і конфіскацію доходів, одержаних злочинним шляхом, і про фінансування тероризму (CETS № 198), Марокко, яке також було запрошене, є єдиною неєвропейською державою, що ратифікувала її. Хоча Російська Федерація припинила членство в РЄ у 2022 році, вона залишається учасницею кількох конвенцій, до яких не подала денонсації, на відміну від ЄСПЛ [90].

Двома основними причинами для розробки договору Ради Європи про Штучний інтелект полягали в тому, щоб мати спільний текст для всіх європейських держав, включаючи Велику Британію після Brexit, і взяти участь у глобальній гонці за те, щоб першим прийняти регламент щодо Штучного інтелекту, в надії слугувати моделлю принаймні для плюралістичних демократій [90].

Додатковою причиною, очевидно, була спроба запропонувати текст, який міг би прийняти велика кількість інших держав, включаючи Сполучені Штати Америки. Участь останнього в переговорах, зокрема в останньому засіданні САІ з 11 по 14 березня 2024 року, призвела, однак, до зменшення сфери його застосування, зокрема тому, що врешті-решт було вирішено, що Рамкова конвенція не застосовуватиметься до приватного сектора. Звичайно, це жодним чином не заважає учасникам майбутньої Конвенції приймати більш інклюзивне законодавство, як це буде у випадку з державами-членами ЄС через АІА.

Як зазначалося вище, інструментом, який використовується Радою Європи для регулювання Штучного інтелекту, є рамкова конвенція, тобто міжнародний договір, відомий як Рамкова конвенція про Штучний інтелект, права людини, демократію та верховенство права, тоді як ЄС використовує регламент, відомий як Регламент Європейського Парламенту та Ради щодо гармонізованих правил у сфері Штучного інтелекту (Artificial Intelligence Act) та внесення змін до деяких законодавчих актів Союзу.

Право РЄ з цієї формальної точки зору: немає юридичної різниці між Конвенцією, Рамковою конвенцією, Угодою, Протоколом, Домовленістю чи навіть Хартією [8], такою як Європейська соціальна хартія, а також Статутом; всі вони є договорами в сенсі міжнародного публічного права. Крім того, існує лише дві офіційні мови Ради Європи: англійська та французька. На початок березня 2024 року із 226 підписаних угод РЄ є три «рамкові» конвенції: Рамкова конвенція про захист національних меншин від 01.02.1995 (ETS № 157) та Рамкова конвенція про цінність культурної спадщини для суспільства

від 27/10/2005 (STCE № 199), до якої можна додати Європейську рамкову конвенцію про транскордонне співробітництво між територіальними громадами або органами влади від 21/05/1980 (ETS № 106) французькою назвою є *Convention-cadre européenne sur la coopération transfrontalière des collectivités ou autorités territoriales*. За словами Генерального директорату Ради Європи з питань демократії та людської гідності, «єдина відмінність між «конвенціями» та «угодами» полягає у формі, у якій держава може висловити свою згоду на обов'язковість для неї. Угоди можуть бути підписані із застереженням щодо ратифікації, прийняття чи затвердження або без нього. Конвенції, в принципі, можуть бути ратифіковані» [64].

Як відомо, відмінності між Договором РЄ та Регламентом ЄС по суті пов'язані з тим фактом, що перший є обов'язковим лише для держав, які його підписали та ратифікували, де це доцільно, тоді як другий є обов'язковим для всього ЄС. Держави-члени, якщо немає виняткових відступів, як правило, тимчасових – або на основі протоколів, що стосуються Ірландії та Данії (і Великобританії до Brexit). Крім того, існують значні відмінності через те, що компетенція ЄС визначена набагато точніше, а тому більш обмежена, ніж компетенція Ради Європи.

На відміну від директив, регламентів і рішень загального характеру ЄС, які в принципі застосовуються безпосередньо до всіх держав-членів, договори Ради Європи застосовуються лише до держав, які їх підписали та ратифікували, якщо цього вимагає їх Конституція.

Якщо не зазначено інше, держави-учасниці можуть робити застереження або заяви до Конвенцій РЄ під час підписання або здачі на зберігання акта про ратифікацію. Об'єктом і наслідком застереження або заяви може бути визначення того, як договір має застосовуватися по відношенню до держави-учасниці. Застереження загального характеру не допускаються щодо ЄКПЛ; застереження може бути зроблено лише щодо окремого положення Конвенції «в тій мірі, в якій закон, що є чинним на її території, не відповідає цьому

положенню». У статті 34 проекту FCIA передбачено «Жодні застереження не можуть бути зроблені щодо будь-якого положення цієї Конвенції» за єдиним винятком, передбаченим у статті 33 щодо федеративних держав, який може бути необхідним для держав, які не є членами Ради Європи, наприклад як Сполучені Штати Америки чи Канада зокрема. Можна припустити, що положення статті 27 – Наслідки Конвенції щодо держав-членів ЄС мають на меті уникнути застережень з боку ЄС або його держав-членів [103].

ЄКПЛ безпосередньо застосовується в більшості сторін. Пряме застосування означає, що на Конвенцію можна посилається в усіх національних судах. Це означає, що інституції відповідної держави – законодавчі, адміністративні та судові органи – зобов'язані поважати Конвенцію, але для особи може бути складніше забезпечити дотримання прав, гарантованих Конвенцією. Випадок FCIA більш делікатний. З одного боку, необхідно буде визначити, наскільки його положення є достатньо точними, щоб вважатися такими, що діють самі по собі, що буде відрізнятися від однієї держави-учасниці до іншої. Зокрема, є держави, де суди самі ухвалюють рішення щодо тлумачення договору, а в інших вони звертаються за тлумаченням до Міністерства закордонних справ. Крім того, є держави, де Конституція чітко вважає договори вищими над законом, і інші, де це питання не вирішено. Це призведе до далекого від однакового застосування положень FCIA, особливо тому, що, на відміну від ЄКПЛ чи Хартії соціальних прав, воно не передбачає створення судового органу, такого як Європейський суд з прав людини чи квазісудовий такий орган, як Європейський комітет із соціальних прав, для вирішення спорів, що виникають у зв'язку з його застосуванням.

Тим не менш, як Суд, так і Комітет посилаються у своїй судовій практиці на всі відповідні інструменти Ради Європи, а також на інші інструменти міжнародного права як на корисний контекст для здійснення своєї юрисдикції

3.3. Глобальна ініціатива з управління Штучним інтелектом

AI Governance Alliance відзначив важливу віху у своїй спільній подорожі з використання трансформаційної сили Штучного інтелекту та генеративного ШІ в різних секторах [120].

Зустріч спільноти AI Governance Alliance, що відбулася 28 травня 2024 року в Сан-Франциско, зібрала 118 лідерів з державного та приватного секторів, академічних кіл і громадянського суспільства для цілого дня діалогу та роботи над просуванням бачення та цілей Альянсу. Він також відзначив захоплюючий і продуктивний рік з моменту заснування Альянсу в червні 2023 року.

Будучи головною ініціативою Всесвітнього економічного форуму та частиною Центру четвертої промислової революції, Альянс спрямований на сприяння відповідальному та етичному розвитку ШІ. Його місія полягає в сприянні прозорості, інклюзивності та підзвітності в Штучному інтелекті, гарантуючи, що технологічний прогрес принесе користь усім зацікавленим сторонам у всьому світі [120].

З моменту заснування Альянс старанно працював над впровадженням найкращих глобальних практик управління ШІ. Його ранні зусилля пов'язані з Рекомендаціями Presidio, які містять вказівки щодо етичного використання та міжнародної співпраці в управлінні ШІ.

У січні 2024 року Альянс випустив першу серію інформаційних документів, у яких об'єднав уроки з його робочих потоків, щоб скерувати зацікавлених сторін до скоординованих стратегій управління Штучним інтелектом, які додатково обговорювалися на зустрічі в Мей.

Основні моменти зустрічі спільноти з управління ШІ включають:

1. Глибоке занурення в інклюзивний ШІ.
2. Дослідження конвергенції ШІ з іншими передовими технологіями.
3. Удосконалення в існуючих робочих групах

На зустрічі також було досягнуто прогресу в трьох основних напрямках роботи Альянсу:

1. Робоча група безпечних систем і технологій: група визначила агентів Штучного інтелекту як системи, що складаються з таких компонентів, як контекстні вікна, пам'ять, дії та автономія, що працюють на основі великих мовних моделей і мультимодальних моделей. Вони класифікували агентів ШІ на основі їхніх функцій, фізичної чи віртуальної форми та автономності. Під час обговорення також було наголошено на вирішенні проблем безпеки та управління, таких як неправильне використання та суспільні розлади. Вони підкреслили важливість чіткої практики управління, включаючи прозорість, відповідність людським цінностям, чесність, яку можна перевірити, і постфактум аудит [120].

2. Робоча група з стійкого управління та регулювання. Ця група підкреслила необхідність розробки прозорих, інклюзивних і підзвітних структур управління Штучним інтелектом шляхом міжнародної співпраці та залучення різноманітних зацікавлених сторін, включаючи різні сектори, малі підприємства та регіони, такі як Глобальний Південь. Обговорення точилося навколо прогалин у чинному законодавстві, ролі у забезпеченні підзвітності Штучного інтелекту, механізмів розкриття інформації та критичної важливості включення різних точок зору для ефективного управління та циклів зворотного зв'язку. Ключові питання, визначені для міжнародного співробітництва, включали гармонізацію стандартів, посилення захисту даних, покращення прозорості та доступу, а також створення стратегічного нагляду та механізмів гнучкого управління для вирішення нових проблем, таких як правила для багатоагентних систем. Робоча група запропонувала створити Global AI Governance Sandbox для тестування та вдосконалення цих структур з метою гармонізації стандартів і посилення глобальної співпраці з питань AI.

3. Робоча група з відповідального застосування та трансформації: Група детально розповіла про трансформаційний потенціал Штучного інтелекту в

різних галузях і про те, як організаціям потрібно прийняти орієнтовану на людину культуру Штучного інтелекту, щоб вижити та процвітати в епоху Штучного інтелекту. Дискусії сприяли необхідності думати в довгостроковій перспективі та прийняти перспективний підхід, оскільки очікується, що ШІ порушить існуючі ланцюжки створення вартості та створить нові бізнес-моделі. Сесія завершилася сильним наголосом на потребі в самоврядуванні галузей промисловості в умовах розвитку технологій і зміни нормативного середовища.

Дивлячись у майбутнє, Альянс управління Штучним інтелектом залишається відданим побудові інклюзивної, відповідальної, стійкої та надійної екосистеми ШІ, яка принесе користь зацікавленим сторонам у всьому світі. Запрошуємо всіх, хто бажає взяти участь у цьому важливому заході під час виконання місії.

Майбутні події включають африканський саміт з питань Штучного інтелекту в Руанді цього жовтня та щорічну зустріч Всесвітнього економічного форуму в Давосі в січні 2025 року. Крім того, на горизонті є кілька публікацій, у тому числі звіти про управління та регулювання, а також про агентів Штучного інтелекту, які мають вийти восени, а потім — публікації про Штучний інтелект та промисловість, а також регіональні плани розвитку Штучного інтелекту в січні [120].

Важливим етапом розвитку ШВ стало започаткування консультативного органу. Генеральний секретар ООН Антоніу Гутерріш оголосив про створення нового консультативного органу з розгляду ризиків і можливостей у міжнародному управлінні Штучним інтелектом (ШІ). Консультативний орган високого рівня зі Штучного інтелекту забезпечить глобальну, міждисциплінарну та багатосторонню дискусію щодо управління ШІ [119].

Виступаючи під час відкриття Консультативного органу високого рівня, Генеральний секретар підкреслив «надзвичайний прогрес» у можливостях ШІ та його використанні за останній рік. Він сказав, що Штучний інтелект може сприяти прогресу людства, активізуючи кліматичні дії та зусилля з досягнення

ЦУР до 2030 року, за умови, що технології Штучного інтелекту будуть використовуватися відповідально та доступні для всіх. Можливості варіюються від «прогнозування та вирішення криз до розгортання послуг охорони здоров'я та освіти» [119]. Країни, що розвиваються, могли б «обійти застарілі технології та надати послуги безпосередньо людям» [119].

«За нинішніх обставин, — попередив Гутерріш, — експертиза ШІ зосереджена в кількох компаніях і країнах», що може посилити глобальну нерівність і цифровий розрив. Він вказав на потенційну шкоду Штучного інтелекту, включаючи дезінформацію та дезінформацію, укорінення упередженості та дискримінації, стеження та вторгнення в приватне життя та шахрайство, серед інших порушень прав людини, і застеріг щодо його зловмисного використання, яке може «підірвати довіру до інститутів, послаблюють соціальну згуртованість і загрожують самій демократії» [119].

Консультативний орган високого рівня вивчить способи зв'язати різні ініціативи з управління Штучним інтелектом і до кінця 2023 року надасть попередні рекомендації щодо того, як використовувати Штучний інтелект для прискорення досягнення ЦСР у таких сферах:

- Міжнародне управління ШІ;
- Спільне розуміння ризиків і викликів;
- Ключові можливості та стимули.

Рекомендації Консультативного органу підтримуватимуть підготовку до Саміту майбутнього 2024 року та сприятимуть переговорам щодо запропонованого Глобального цифрового договору. Організація консультуватиметься з Науковою консультативною радою Генерального секретаря щодо того, як використовувати переваги ШІ для прискорення прогресу «на благо всіх».

Консультативний орган високого рівня є гендерно збалансованим, географічно різноманітним і охоплює покоління. Переносячи досвід між урядом, приватним сектором, технологічною спільнотою, громадянським

суспільством і науковими колами, 39 членів Організації працюватимуть у своїй особистій якості. Вони також проведуть широкі, широкі та прозорі консультації, щоб почути погляди всіх зацікавлених сторін [119].

Глобальні можливості та виклики, створені ШІ, вимагають широкого залучення та участі, тому в умовах сьогодення виникло явище глобального партнерства з ШІ [118].

Глобальне партнерство зі Штучного інтелекту (GPAI) — це інтегроване партнерство, яке об'єднує членів ОЕСР та країни GPAI для просування амбітної програми впровадження орієнтованого на людину безпечного, захищеного та надійного Штучного інтелекту (ШІ), втіленого в принципах Рекомендації ОЕСР. на ШІ [118].

Інтегроване партнерство об'єднує країни на рівних для використання синергії та взаємодоповнюваності між GPAI та робочою програмою ОЕСР щодо Штучного інтелекту, сприяння всеохоплюючій участі як на рівні управління, так і на експертному рівні, а також сприяння більш ефективним процесам, зменшуючи витрати та дублювання [118].

На фоні створених організацій та об'єднань виникає питання етики використання ШІ.

У листопаді 2021 року ЮНЕСКО випустило перший глобальний стандарт етики Штучного інтелекту – «Рекомендацію щодо етики Штучного інтелекту». Цю структуру прийняли всі 193 держави-члени.

«Захист прав і гідності людини є наріжним каменем Рекомендації, яка базується на просуванні фундаментальних принципів, таких як прозорість і справедливість, завжди пам'ятаючи про важливість людського контролю за системами ШІ. Проте те, що робить Рекомендацію винятково придатною, так це її широкі сфери політичних дій, які дозволяють політикам втілювати основні цінності та принципи в дію щодо управління даними, навколишнього середовища та екосистем, гендерних питань, освіти та досліджень, а також

охорони здоров'я та соціального благополуччя, серед багато інших сфер» - наголошує ЮНЕСКО [117].

Всі наведені процеси дають сформувати уявлення про розвиток управління Штучним інтелектом та його подальше просування.

ВИСНОВКИ

Термін «Штучний інтелект» має широке і вузьке значення. У вузькому сенсі це інформація, що створюється автоматизованими системами. Загалом, ШІ означає здатність системи виконувати людські функції, аналізуючи фактори та приймаючи рішення на основі досвіду. Джон Маккарті вперше визначив ШІ в 1956 році як створення агентів, які адаптуються до нових ситуацій і досягають цілей, що недоступні традиційним системам. High-Level Expert Group on Artificial Intelligence Європейської Комісії визначає ШІ як систему, що діє в фізичному або цифровому світі, аналізуючи дані для досягнення цілей. Згідно з Joint Research Centre, ШІ - це система, що розпізнає своє оточення і дані для вибору найкращого способу дій.

Ризики Штучного інтелекту включають кілька основних загроз. По-перше, зловмисники можуть використовувати генеративний ШІ для масштабування атак, вдосконалення методів програм-вимагачів і фішингових атак, вони можуть маніпулювати автономними системами, такими як безпілотні автомобілі або медичні системи, що може загрожувати фізичній безпеці. По-друге, автоматизоване шкідливе програмне забезпечення може створюватися за допомогою ШІ, такого як ChatGPT, що може витіснити розробників програмного забезпечення. Ризики конфіденційності включають витік особистої інформації через помилки або атаки на системи ШІ. Також існує загроза крадіжки і маніпуляції моделями ШІ, що може посилити зловмисні дії. Інша проблема - маніпуляція даними, що може призвести до небажаних або шкідливих результатів. Уособлення, створене за допомогою ШІ, може бути використане для віртуальних шахрайств і фальшивих записів.

Серед переваг ШІ може проактивно запобігати кібератакам, розпізнаючи та нейтралізуючи загрози, ще до того, як вони завдадуть шкоди. Він допомагає покращити реагування на інциденти, швидко аналізуючи дані та автоматично усуваючи загрози в режимі реального часу. Також він підвищує точність

виявлення загроз, зменшуючи людські помилки та звільняючи час для більш стратегічної роботи. ШІ ефективно обробляє великі обсяги даних, виявляючи ненормальні моделі та проблеми у мережі, підвищує ефективність, спрощуючи розробку політики і процедур, вдосконалюють навчання з кібербезпеки, створюючи реалістичні симуляції кібератак, забезпечує масштабованість для росту організації та адаптації до змін у кіберризиках. Автоматизоване тестування і моніторинг вразливостей забезпечують безперервну перевірку систем безпеки. Штучний інтелект також використовується для прогнозування ризиків, аутентифікації користувачів і виявлення шахрайства.

Штучний інтелект має величезний потенціал для покращення суспільства, але також створює серйозні виклики для прав людини, демократії та верховенства права. Його вплив є як позитивний, так і негативний.

ШІ може бути використаний для передбачення злочинів або визначення рівня ризику повторного злочину, що впливає на право на свободу і безпеку. Також він активно використовується для масового стеження, розпізнавання обличчя, і обробки біометричних даних, що ставить під загрозу право на приватність (ст. 8 ЄКПЛ). ШІ може закріплювати або посилювати дискримінаційні упередження, як це було продемонстровано на прикладі рекрутингової системи Amazon, що віддавала перевагу чоловікам через упереджені дані. ШІ, що використовується в системах рекомендацій, новинних агрегаторах та пошукових системах, може обмежувати доступ до інформації або надавати персоналізовану інформацію, що упереджена чи фальсифікована. Це впливає на демократичний дискурс, дозволяючи створювати «інформаційні ехо-камери» та маніпулювати громадською думкою.

Персоналізована реклама і мікротаргетинг можуть маніпулювати виборцями шляхом доставки специфічних повідомлень різним групам, що підриває чесність і прозорість виборчого процесу. ШІ може бути використано для створення глибоких фейків і дезінформації, що ускладнює ідентифікацію правдивої інформації і підриває довіру до медіа та демократичного процесу.

Автоматизація рішень у важливих сферах, таких як фінанси, енергетика, транспорт, може створювати нові системні ризики і загрози безпеці суспільних систем. ШІ може погіршити умови праці, порушувати право на гідні умови праці та рівне ставлення, особливо коли воно використовується для моніторингу працівників або оцінки їх продуктивності.

Для врегулювання наведених проблем необхідно звернутися до організаційно-правового регулювання. Країни Європейського Союзу, США та Великобританія вже запровадили низку ініціатив та правових актів, спрямованих на регулювання впровадження Штучного інтелекту (ШІ) для мінімізації ризиків та підвищення прозорості його використання. ЄС, зокрема, прийняв Регламент про ШІ (Artificial Intelligence Act), що встановлює гармонізовані правила щодо ризиків та етичного використання технологій. У США діють рекомендації Національного інституту стандартів і технологій (NIST), які акцентують на управлінні ризиками та відповідальному впровадженні ШІ. У Великобританії проводиться адаптація існуючих законодавчих норм для врахування особливостей цифрових технологій, включаючи вплив ШІ на права працівників та кібербезпеку.

Ключовим елементом цих ініціатив є інтеграція принципів етики, прозорості та захисту прав людини, що дозволяє країнам ефективно управляти викликами, які виникають через автоматизацію. Такі підходи допомагають забезпечити безпечне використання ШІ в секторах, які мають критичне значення для суспільства, включаючи фінанси, енергетику та транспорт. Водночас ці ініціативи сприяють міжнародній співпраці для створення глобальних стандартів регулювання технологій ШІ, що дозволяє уникнути фрагментації в підходах і забезпечити збалансовану інтеграцію інновацій у різних країнах.

В Україні у 2020 році була створена Концепція розвитку Штучного інтелекту, яка визначає мету, принципи та завдання технологій ШІ. Національна асоціація юристів України сформувала Робочу групу для аналізу

правових питань ШІ, включаючи захист персональних даних. У грудні 2022 року внесено зміни до закону про авторське право, що стосуються прав на твори, створені ШІ, запроваджено права *sui generis*. У 2024 році Міністерство цифрової трансформації розробило White Paper, що описує підхід до регулювання ШІ. Документ має допомогти компаніям підготуватися до нових вимог та забезпечити безпеку продуктів. Розроблено рекомендації для бізнесу, кодекси поведінки та регуляторну пісочницю.

Генеральна Асамблея ООН 11 березня 2024 року ухвалила резолюцію, яка регулює розробку, виробництво та використання Штучного інтелекту (ШІ). Європейський парламент також прийняв Регламент про ШІ в березні 2024 року.

Спеціальний комітет Ради Європи зі Штучного інтелекту (CAI) розробив проект Конвенції про Штучний інтелект (Конвенція про ШІ), що встановлює загальні принципи, включаючи людську гідність, прозорість, підзвітність і захист приватного життя та персональних даних. Конвенція також підкреслює важливість безпечних інновацій, надійності та рівних можливостей. Глава IV присвячена засобам правового захисту та процесуальним гарантіям, а Глава VI зосереджена на впровадженні та просуванні цифрової грамотності та навичок. Конвенція має на меті сприяти інноваціям, забезпечуючи при цьому повагу до прав людини та верховенства права. Комісія ЄС брала участь у переговорах, і Конвенція була прийнята в березні 2024 року.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аналіз національних стратегій розвитку Штучного інтелекту. URL: https://ippi.org.ua/sites/default/files/9_24.pdf (дата звернення 21.08.2024).
2. Баранов О.А. Визначення терміну «Штучний інтелект». Інформація і право. 2023. № 1 (44). С. 32-49. URL: https://ippi.org.ua/sites/default/files/5_28.pdf (дата звернення 17.07.2024).
3. Баранов. О.А. визначення терміну “штучний інтелект”. інформація і право. № 1(44). 2023. URL: https://ippi.org.ua/sites/default/files/5_28.pdf (дата звернення 17.04.2024).
4. Висновок 20/2022 Європейського інспектора із захисту даних щодо Рекомендації щодо Рішення Ради про дозвіл на початок переговорів від імені Європейського Союзу з метою укладення Конвенції Ради Європи про Штучний інтелект, права людини, демократію та верховенство права. URL: https://www.edps.europa.eu/system/files/2022-10/22-10-13_edps-opinion-ai-human-rights-democracy-rule-of-law_en.pdf. (дата звернення: 20.07.2024).
5. Вінникова Н. А. Штучний інтелект як чинник геополітичного впливу. Наукова періодика Каразінського університету. 2021. Вип. 40. С 3-4. URL: <https://periodicals.karazin.ua/politology/article/view/18660/16983> (дата звернення: 10.07.2024).
6. Вінникова Н.А. Штучний інтелект у контексті глобального управління. Політичні проблеми міжнародних систем та глобального розвитку. Вип 3. 2022. URL: http://politicus.od.ua/3_2022/10.pdf (дата звернення 21.08.2024).
7. Габер Є. Міжнародні відносини в епоху Штучного інтелекту: чи мають майбутнє дипломатичні перекладачі?. Дзеркало тижня | Mirror Weekly. 2019. С. 2-3. URL: https://zn.ua/ukr/international/mizhнародni-vidnosini-vepohu-shtuchnogo-intelektu-chi-mayut-maybutnye-diplomatichni-perekladachi-326099_.html (дата звернення: 10.07.2024).

8. Етична хартія закріплює п'ять принципів щодо використання штучного інтелекту під час здійснення правосуддя. URL: <https://cedem.org.ua/analytics/shtuchnyj-intelekt-pravosuddia/> (дата звернення 19.11.2024).

9. Європейська конвенція про захист прав людини. URL: https://www.google.com/search?q=%D0%84%D0%9A%D0%9F%D0%A7&oq=%D0%84%D0%9A%D0%9F%D0%A7&gs_lcrp=EgZjaHJvbWUyBggAEEUYOТИКCAEQABiABBiiBDIKCAIQABiABBiiBDIKCAMQABiABBiiBDIKCAQQABiABBiiBDIKCAUQABiABBiiBNIBBzU4OGowajmoAgCwAgE&sourceid=chrome&ie=UTF-8#:~:text=%D0%84%D0%B2%D1%80%D0%BE%D0%BF%D0%B5%D0%B9%D1%81%D1%8C%D0%BA%D0%B0%20%D0%BA%D0%BE%D0%BD%D0%B2%D0%B5%D0%BD%D1%86%D1%96%D1%8F%20%D0%BF%D1%80%D0%BE%20%D0%B7%D0%B0%D1%85%D0%B8%D1%81%D1%82%20%D0%BF%D1%80%D0%B0%D0%B2%20%D0%BB%D1%8E%D0%B4%D0%B8%D0%BD%D0%B8 (дата звернення 19.11.2024).

10. Загальний регламент про захист даних. URL: <https://gdpr-text.com/uk/> (дата звернення 19.11.2024).

11. Застосування ШІ на роботі та в побуті. URL: <https://www.softkey.ua/ua/useful/articles/zastosuvannya-shi-na-roboti-ta-v-pobuti/> (дата звернення 21.08.2024).

12. Застосування Штучного Інтелекту у Повсякденному Житті. URL: https://futurenow.com.ua/5-prykladiv-shtuchnogo-intelektu-u-shhodennomu-zhytti/#google_vignette (дата звернення: 26.06.2024).

13. ЗУ «Про схвалення Концепції розвитку штучного інтелекту в Україні» URL: <https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%80#Text> (дата звернення 19.11.2024).

14. І знову про Штучний інтелект. Допомога, загроза чи пусті балачки?
URL: <https://yur-gazeta.com/publications/practice/inshe/i-znovu-pro-shtuchniy-intelekt-dopomoga-zagroza-chi-pusti-balachki.html> (дата звернення: 24.07.2024).

15. Кулеба Д. Дипломатія майбутнього. Українська правда – Блоги.
URL: <https://blogs.pravda.com.ua/authors/kuleba/5a3cc08726814/> (дата звернення: 10.07.2024).

16. Місто-держава Ватикан є суверенним органом згідно з міжнародним публічним правом, відмінним від Святого Престолу (тобто глави Римо-Католицької Церкви), і користується загальним визнанням, але воно має особливий характер, який пояснює тим, що, крім Організації, в яких Святий Престол бере участь як постійний спостерігач, такі як Рада Європи, місто-держава Ватикан є членом лише кількох міжнародних організацій, таких як Всесвітній поштовий союз (ВПС), Міжнародний союз телекомунікацій (МСЕ), Міжнародне агентство з атомної енергії (МАГАТЕ) та Всесвітня туристична організація (ЮНВТО). V. URL: <https://www.vaticanstate.va/it/stato-governo/note-general/origini-natura.html> (дата звернення: 20.07.2024).

17. Муравська Юлія. Сліпченко Тетяна. Правове регулювання Штучного інтелекту в Україні та світі. Актуальні проблеми правознавства. 2024. № 1 (37). С. 188-195. URL: <https://appj.wunu.edu.ua/index.php/apl/article/view/1788/1842> (дата звернення: 28.07.2024).

18. Новини РЄ від 12 жовтня 2023 р. «Генеральний секретар Марія Пейчинович Бурич зустрілася з єврокомісаром з питань юстиції Дідьє Рейндерсом. Зустріч була зосереджена на співпраці між Радою Європи та Європейським Союзом, а також на поточній підготовці до Конвенції про Штучний інтелект». URL: <https://www.coe.int/en/web/portal/-/secretary-general-meets-european-commissioner-for-justice> (дата звернення: 10.07.2024).

19. Проект пояснювальної доповіді. URL: <https://rm.coe.int/-1497-10-1b-committee-on-artificial-intelligence-cai-b-draft-framework-convention/1680af0734&format=native> (дата звернення: 10.07.2024).

20. Проект резолюції Генеральної Асамблеї Організації Об'єднаних Націй від 11 березня 2024 року щодо використання можливостей безпечних, захищених і надійних систем Штучного інтелекту для сталого розвитку. URL: <https://digitallibrary.un.org/record/4040897?ln=en&v=pdf> (дата звернення: 28.07.2024).

21. Регулювання Штучного інтелекту в Україні: Міністерство цифрової трансформації презентувало White Paper. URL: <https://www.kmu.gov.ua/en/news/rehuliuvannia-shtuchnoho-intelektu-v-ukraini-mintsyfry-prezentuie-bilu-knyhu> (дата звернення: 28.07.2024).

22. Рішення Ради (ЄС) 2022/2349 від 21 листопада 2022 року про дозвіл на початок переговорів від імені Європейського Союзу щодо Конвенції Ради Європи про Штучний інтелект, права людини, демократію та верховенство права. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022D2349> (дата звернення: 10.07.2024).

23. Слюсар В. І. Штучний інтелект у сфері безпеки та оборони. Щодо проєкту Стратегії розвитку Штучного Інтелекту в Україні НА 2022–2030 рр. 2020. С. 132–157. URL: https://www.slyusar.kiev.ua/Slyusar_AI_2022-1-1_ua.pdf (дата звернення: 10.07.2024).

24. Штучний інтелект: тренди, перспективи, ризики. URL: <https://svit.kpi.ua/2024/06/19/%D1%88%D1%82%D1%83%D1%87%D0%BD%D0%B8%D0%B9-%D1%96%D0%BD%D1%82%D0%B5%D0%BB%D0%B5%D0%BA%D1%82-%D1%82%D1%80%D0%B5%D0%BD%D0%B4%D0%B8-%D0%BF%D0%B5%D1%80%D1%81%D0%BF%D0%B5%D0%BA%D1%82%D0%B8%D0%B2%D0%B8-2/> (дата звернення: 24.07.2024).

25. Юшкевич О. Є. Роль Штучного інтелекту в прийнятті зовнішньополітичних рішень (на прикладі США). Репозитарій Національного Авіаційного Університету: Home. 2022. URL: https://dspace.nau.edu.ua/bitstream/NAU/55611/1/ФМВ_291_МІ_Юшкевич_О..pdf (дата звернення 21.08.2024).

26. 18 квітня Парламентська асамблея Ради Європи одноголосно схвалила проект «Конвенції про Штучний інтелект, права людини, демократію та верховенство права», хоча й висловила жаль щодо дисбалансу між державним та приватним секторами проекту. URL: <https://pace.coe.int/en/verbatim/2024-04-18/pm/en> (дата звернення 21.08.2024).

27. 4.6: Рогеріан Аргумент. URL: https://ukrayinska.libretexts.org/%D0%93%D1%83%D0%BC%D0%B0%D0%BD%D1%96%D1%82%D0%B0%D1%80%D0%BD%D1%96_%B5%D0%BD%D1%82 (дата звернення: 10.09.2024).

28. Abduljabbar, R. et al. “Applications of Artificial Intelligence in Transport: An Overview”. Computer Science. Sustainability: n. pag. 2019. URL: <https://www.mdpi.com/2071-1050/11/1/189> (дата звернення 21.08.2024).

29. AI and cybersecurity: How to navigate the risks and opportunities. URL: <https://www.weforum.org/agenda/2024/02/ai-cybersecurity-how-to-navigate-the-risks-and-opportunities/> (дата звернення 21.08.2024).

30. AI in Cyber Security: Pros and Cons, and What it Means for Your Business URL: <https://www.terranosecurity.com/blog/ai-in-cyber-security> (дата звернення 21.08.2024).

31. AI in Cyber Security: Risks of AI. URL: <https://www.malwarebytes.com/cybersecurity/basics/risks-of-ai-in-cyber-security> (дата звернення 21.08.2024).

32. AI in Cybersecurity: Upsides, Downsides, Opportunities, and Dangers URL: <https://www.logicgate.com/blog/ai-in-cybersecurity-upsides-downsides-opportunities-and-dangers/> (дата звернення 21.08.2024).

33. Altus Ashley. AI and copyright: Protecting ownership and intellectual property rights. 07/10/2024. URL: <https://outshift.cisco.com/blog/ai-copyright-protecting-ownership-intellectual-property> (дата звернення 19.11.2024).

34. Artificial Intelligence Act. URL: MEPs adopt landmark law: <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law> (дата звернення 21.08.2024).

35. Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI. URL: <https://www.europarl.europa.eu/news/en/pressroom/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai> (дата звернення 21.08.2024).

36. Benizri Itsiq, Evers Arianna, Togawa Mercer Shannon, Jessani Ali A. A Comparative Perspective on AI Regulation. LAWFARE. July 17. 2023. URL: <https://www.lawfaremedia.org/article/a-comparative-perspective-on-ai-regulation> (дата звернення: 28.07.2024).

37. Berry Kathy. UK fails to agree AI/copyright code of practice. Linklaters. 8 February 2024. URL: <https://www.linklaters.com/insights/blogs/digilinks/2024/february/uk-fails-to-agree-ai---copyright-code-of-practice> (дата звернення 19.11.2024).

38. Beyer, H. et al. «olving conservation planning problems with integer linear programming». Ecological Modelling 328. 2016. 14-22. URL: <https://www.sciencedirect.com/science/article/pii/S0304380016300217> (дата звернення: 28.07.2024).

39. Broersma Matthew. China Regulator Warns AI Poses Challenge To Society. SILICON. June 26. 2023. URL: <https://www.silicon.co.uk/e-innovation/artificial-intelligence/china-ai-regulation-517977> (дата звернення 19.11.2024).

40. Čerka, Paulius, Jurgita Grigienė and Gintarė Sirbikytė. “Liability for damages caused by artificial intelligence”. Comput. Law Secur. Rev. 31. 2015. C.

376-389. URL:
<https://www.sciencedirect.com/science/article/pii/S026736491500062X> (дата
 звернення: 28.07.2024).

41. Chatila, Raja. Ethically Aligned Design. For Public Discussion 136.
 2016. URL: [https://standards.ieee.org/wp-
 content/uploads/import/documents/other/ead_v2.pdf](https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf) (дата звернення: 28.07.2024).

42. Chen, Jim Q. and Wingfield, Thomas. "Human-Machine Teaming and
 Its Legal and Ethical Implications". Military Cyber Affairs 4(2):2. 2020. URL:
<https://digitalcommons.usf.edu/mca/vol4/iss2/2/> (дата звернення 21.08.2024).

43. Chien, Steve and Robert Morris. "Space Applications of Artificial
 Intelligence." AI Mag. 35. 2014. C. 3-6. URL:
<https://onlinelibrary.wiley.com/doi/10.1609/aimag.v35i4.2551> (дата звернення:
 28.07.2024).

44. Comité sur l'intelligence artificielle (CAI). URL: [https://rm.coe.int/-
 1493-10-1b-committee-on-artificial-intelligence-cai-b-draft-framework/1680aee411](https://rm.coe.int/-1493-10-1b-committee-on-artificial-intelligence-cai-b-draft-framework/1680aee411) /
[https://rm.coe.int/-1493-10-
 1b-comite-sur-l-intelligence-artificielle-cai-b-projet-de-
 con/1680aee410](https://rm.coe.int/-1493-10-1b-comite-sur-l-intelligence-artificielle-cai-b-projet-de-con/1680aee410) (дата звернення: 15.07.2024).

45. Committee on artificial intelligence (CAI). URL: [https://rm.coe.int/cai-
 2023-28-draft-framework-convention/1680ade043](https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043). (дата звернення: 28.07.2024).

46. Copyright and Artificial Intelligence Part 1: Digital Replicas. A
 REPORT of the register of copyrights july 2024. URL:
[https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-
 Replicas-Report.pdf](https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf) (дата звернення 19.11.2024).

47. Cordell Neville, Potts Beverley, Dickens Robert. Ownership of AI-
 generated content in the UK. A&O Shearman. Aug 20. 2024. URL:
[https://www.aoshearman.com/en/insights/ownership-of-ai-generated-content-in-the-
 uk](https://www.aoshearman.com/en/insights/ownership-of-ai-generated-content-in-the-uk) (дата звернення 19.11.2024).

48. Council of Europe and Artificial Intelligence. URL: www.coe.int/ai .
 (дата звернення: 28.07.2024).

49. Critical infrastructure URL: <https://dig.watch/topics/critical-infrastructure> (дата звернення: 28.07.2024).

50. Dennis Gill. UK government AI report confirms decision on protection of copyright works. Pinsent Masons. 19 Jan 2024. URL: <https://www.pinsentmasons.com/out-law/news/ai-report-confirms-decision-on-protection-of-copyright-works> (дата звернення 19.11.2024).

51. Devedzic, V. "Identity of AI". Discov Artif Intell 2. 2022. URL: <https://link.springer.com/article/10.1007/s44163-023-00054-8> (дата звернення: 28.07.2024).

52. E. Faletti, L'Artificial Intelligence Act Proposal e la regolamentazione degli algoritmi predittivi: luci e ombre. CERIDAP, 4. 2023. 173-217 URL: <https://ceridap.eu/lartificial-intelligence-act-proposal-e-la-regolamentazione-degli-algoritmi-predittivi-luci-e-ombre/?lng=en>. (дата звернення 21.08.2024).

53. Elkus A. The futility of (narrow) speculation about machines and jobs. Medium. URL: <https://medium.com/@Aelkus/the-futility-of-narrow-speculation-about-machines-and-jobs-a116a0672659> (дата звернення 21.08.2024).

54. Engler Alex. The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment. BROOKINGS. April 25. 2023. URL: <https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/> (дата звернення: 28.07.2024).

55. EU AI Act: first regulation on artificial intelligence. The use of artificial intelligence in the EU will be regulated by the AI Act, the world's first comprehensive AI law. European Parliament. 18-06-2024. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> (дата звернення 19.11.2024).

56. EU4DigitalUA and Ukrainian officials present Artificial Intelligence guidelines. URL: <https://eufordigital.eu/eu4digitalua-and-ukrainian-officials-present-artificial-intelligence-guidelines/> (дата звернення: 28.07.2024).

57. European Innovation Council and SMEs Executive Agency. Artificial intelligence and copyright: use of generative AI tools to develop new content. European Commission. 16 July 2024. URL: https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/artificial-intelligence-and-copyright-use-generative-ai-tools-develop-new-content-2024-07-16-0_en (дата звернення 19.11.2024).

58. European Commission, White Paper on Artificial Intelligence - A European approach to excellence and trust, COM(2020) 65 final, 2020. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206#footnote3> (дата звернення 19.11.2024).

59. European Commission. A definition of Artificial Intelligence: main capabilities and scientific disciplines. Report High-Level Expert Group on Artificial Intelligence. 18 December 2018. URL: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> (дата звернення: 28.07.2024).

60. European Parliament. Artificial Intelligence Act. P9_TA(2024)0138 European Parliament legislative resolution of 13 March 2024. URL: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf (дата звернення 19.11.2024).

61. Executive Order On Advancing Racial Equity and Support. URL: <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/> (дата звернення 19.11.2024).

62. Galeotti M. The age of AI diplomacy. The spectator. 2023. URL: <https://www.spectator.co.uk/article/the-ageof-ai-diplomacy/> (date of appeal: 10.04.2024).

63. General Assembly adopts landmark resolution on artificial intelligence. URL: <https://news.un.org/en/story/2024/03/114783> (дата звернення: 27.07.2024).

64. Glossary. URL: <https://www.coe.int/en/web/conventions/glossary> (дата звернення: 28.07.2024).

65. Godefroy James. How Does Artificial Intelligence Affect Intellectual Property Protection? ROUSE. 07 Mar 2024. URL: <https://rouse.com/insights/news/2024/how-does-artificial-intelligence-affect-intellectual-property-protection#:~:text=AI%20technologies%20can%20help%20to,of%20AI%20on%20IP%20rightsholders>. (дата звернення: 28.07.2024).

66. Guadamuz Andres. Artificial intelligence and copyright. WIPO. October 2017. URL: https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html (дата звернення 19.11.2024).

67. Guadamuz Andres - What's the status of AI and copyright?. URL: https://www.linkedin.com/posts/andres-guadamuz_Whats-the-status-of-ai-and-copyright-activity-7210299774544048128-SAxx3 (дата звернення 19.11.2024).

68. H.R.2575 - AI in Government Act of 2020. 116th Congress (2019-2020). URL: <https://www.congress.gov/bill/116th-congress/house-bill/2575> (дата звернення 19.11.2024).

69. Heart on My Sleeve. URL: https://lyrsense.com/olly_murs/heart_on_my_sleeve (дата звернення 19.11.2024).

70. Hickman Tim, Lorenz Sylvia, Teetzmann Constantin, Jha Aishwarya. Long awaited EU AI Act becomes law after publication in the EU's Official Journal. White&Case. 16 July 2024. <https://www.whitecase.com/insight-alert/long-awaited-eu-ai-act-becomes-law-after-publication-eus-official-journal> (дата звернення 19.11.2024).

71. Holzinger A. Causability and explainability of artificial intelligence in medicine. WIREs Data Mining Knowl Discov. 2019. C. 9. URL:

<https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1312> (дата звернення: 28.07.2024).

72. Horowitz M., Mahoney C. Artificial intelligence and the military: technology is only half the battle. War on the Rocks. December 25. 2018. URL: <https://warontherocks.com/2018/12/artificial-intelligence-and-the-military-technology-is-only-half-thebattle/> (дата звернення: 28.07.2024).

73. Interinstitutional File. URL: 2021/0106(COD): <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf>. (дата звернення: 17.07.2024).

74. Interinstitutional File. URL: <https://rm.coe.int/-1497-10-1b-comite-sur-l-intelligence-artificielle-cai-b-projet-de-convention/1680af0733&format=native> (дата звернення: 17.07.2024).

75. Ivy Patrick. How AI May Impact Intellectual Property Rights. Newburn Law. July 11, 2023. URL: <https://www.newburnlaw.com/how-ai-may-impact-intellectual-property-rights> (дата звернення: 28.07.2024).

76. Kanade, Vijay. What Is Artificial Intelligence (AI)? Definition, Types, Goals, Challenges, and Trends in 2022. SpiceWorld 2022. URL: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ai> (дата звернення: 24.04.2024).

77. L. Cotino Hueso, El Convenio sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho del Consejo de Europa, in Revista Administración & Ciudadanía, EGAP, 2024. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4942383 (дата звернення: 28.07.2024).

78. L. Cotino Hueso. El v Aragonský věstník veřejné správy, Monografie XIV, str. 60. 2013. URL: https://scholar.google.com.ua/citations?view_op=search_authors&mauthors=L.+Cotino+Hueso.&hl=uk&oi=ao (дата звернення: 28.07.2024).

79. Lawrence Christie, Cui Isaac, and Ho Daniel E., Implementation Challenges to Three Pillars of America's AI Strategy (Stanford: Stanford RegLab and Stanford University Center for Human-Centered AI, 2022). URL: https://dho.stanford.edu/wp-content/uploads/AI_Implementation.pdf (дата звернення 19.11.2024).

80. LEI Convenio sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho del Consejo de Europa. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4942383 (дата звернення: 28.07.2024).

81. Lenardon, J. The regulation of artificial intelligence. Tilburg University. AI & SOCIETY. 2023. URL: https://www.researchgate.net/publication/369758616_The_regulation_of_artificial_intelligence (дата звернення: 28.07.2024).

82. Li Sophie. Hou Sophia. Enhancing Copyright Protection: Features of China's AI Intelligent Review Tool. ROUSE. 23 Oct 2024. URL: <https://rouse.com/insights/news/2024/enhancing-copyright-protection-features-of-china-s-ai-intelligent-review-tool> (дата звернення 19.11.2024).

83. Maintaining American Leadership in Artificial Intelligence. Executive Order 13859 of February 11, 2019. URL: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence> (дата звернення 19.11.2024).

84. Mission AI: The New System Technology URL: <https://link.springer.com/book/10.1007/978-3-031-21448-6> (дата звернення: 26.07.2024).

85. National Artificial Intelligence Initiative (NAII) Act of 2020 (Division E of the National Defense Authorization Act for Fiscal Year 2021). URL: <https://www.govinfo.gov/content/pkg/PLAW-116publ283/html/PLAW-116publ283.htm> (дата звернення 19.11.2024).

86. OECD. Artificial Intelligence in Society. Paris. OECD Publishing. 2019. 152 p. URL: <https://doi.org/10.1787/eedfee77-en>. (дата звернення: 24.08.2024).

87. Petaković Miloš. Will Artificial Intelligence Kill Copyright? GECICLAW. 10 Apr. 2023. URL: <https://www.geciclaw.com/artificial-intelligence-intellectual-property/> (дата звернення 19.11.2024).

88. Presentation of the Council of Europe's activities on Artificial Intelligence (AI) during the OECD - African Union AI Dialogue. URL: <https://www.coe.int/en/web/artificial-intelligence/-/presentation-of-the-council-of-europe-s-activities-on-artificial-intelligence-ai-during-the-oecd-african-union-ai-dialogue> , (дата звернення: 28.07.2024).

89. Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. A Presidential Document by the Executive Office of the President on 12/08/2020. URL: <https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government> (дата звернення 19.11.2024).

90. Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts: <https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX:52021PC0206> (дата звернення: 17.07.2024).

91. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> (дата звернення 19.11.2024).

92. Racial Equity and Supporting Underserved Communities Through the Federal Government" EO 14091 URL: <https://govinfo.gov/app/details/CFR-2024-title3-vol1/CFR-2024-title3-vol1-eo14091/summary> (дата звернення 19.11.2024).

93. Regulation - 2016/679 - EN - gdpr - EUR-Lex. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (дата звернення 19.11.2024).

94. Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. A Presidential Document by the Executive Office of the President on 11/01/2023. URL: <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence> (дата звернення 19.11.2024).

95. Samoli, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F. and Delipetrev, B., AI WATCH. Defining Artificial Intelligence. Publications Office of the European Union, Luxembourg, 2020. URL: 10.2760/382730 (дата звернення: 02.08.2024).

96. Scharre, P. and M. Horowitz. "Artificial intelligence". What Every Policymaker Needs to Know. Report of the Center for a New American Security's. P.19. 2018. URL: <https://tnsr.org/wp-content/uploads/2020/06/TNSR-AI-Roundtable-PDF.pdf> (дата звернення: 26.07.2024).

97. Scott B., Heumann S., Lorenz P. Artificial intelligence and foreign policy. Stiftung Neue Verantwortung (SNV). 2017. URL: https://www.stiftung-nv.de/sites/default/files/ai_foreign_policy.pdf (дата звернення: 26.07.2024).).

98. Shakey : An Experiment in Robot Planning and Learning. URL: <https://exhibits.stanford.edu/ai/catalog/tt618qz3245> (дата звернення: 02.08.2024).

99. Staff Coursera. The History of AI: A Timeline of Artificial Intelligence. COURSERA. Updated on Oct 25, 2024. URL: <https://www.coursera.org/articles/history-of-ai> (дата звернення: 26.06.2024).

100. Terms of reference of the Committee on Artificial Intelligence (CAI). URL: <https://rm.coe.int/terms-of-reference-of-the-committee-on-artificial-intelligence-cai-/1680ade00f> . (дата звернення: 28.07.2024).

101. The Artificial Intelligence Act - Regulation (EU) 2024/1689 URL: <https://www.artificial-intelligence-act.com/> (дата звернення 19.11.2024).

102. The Association for the Advancement of Artificial Intelligence. URL: <https://aaai.org/> (дата звернення: 07.09.2024).

103. The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments. URL: <https://ceridap.eu/the-council-of-europe-framework-convention-on-artificial-intelligence-vs-the-eu-regulation-two-quite-different-legal-instruments/?lng=en> (дата звернення: 28.07.2024).

104. The Impact of AI on Intellectual Property Practices: Revolutionizing The Legal Profession. URL: <https://www.linkedin.com/pulse/impact-ai-intellectual-property-practices-legal-maurice-bretzfeld> (дата звернення: 09.08.2024).

105. The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law URL: <https://allai.nl/wp-content/uploads/2020/06/The-Impact-of-AI-on-Human-Rights-Democracy-and-the-Rule-of-Law-draft.pdf32f860076e49> (дата звернення: 07.10.2024).

106. The Law Dictionary. Your Free Online Legal Dictionary. Featuring Black's Law Dictionary, 2nd Ed. 2023. «Artificial intelligence». URL: <https://thelawdictionary.org/?s=artificial+intelligence> (дата звернення: 24.09.2024).

107. The rise of AI fake news is creating a 'misinformation superspreader'. URL: <https://www.washingtonpost.com/technology/2023/12/17/ai-fake-news-misinformation/> (дата звернення: 07.09.2024).

108. Tracking developments in artificial intelligence research: constructing and applying a new search strategy. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8550099/> (дата звернення: 07.09.2024).

109. Tsuiman Aleksander. What are the Executive Orders, and how do they impact AI? VERIFF. February 14. 2024. URL: <https://www.veriff.com/fraud/learn/ai-executive-order> (дата звернення 19.11.2024).

110. U.S. Artificial Intelligence Safety Institute. URL: <https://www.nist.gov/aisi> (дата звернення: 28.07.2024).
111. Vallance Patrick. Pro-innovation Regulation of Technologies Review. Digital Technologies. March 2023. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1142883/Pro-innovation_Regulation_of_Technologies_Review_-_Digital_Technologies_report.pdf (дата звернення 19.11.2024).
112. Voelsen D., Stanzel V. Diplomacy and artificial intelligence. Stiftung Wissenschaft und Politik (SWP). 2016. URL: <https://www.swp-berlin.org/en/publication/diplomacy-and-artificial-intelligence#hd-d22571e579> (дата звернення: 28.07.2024).
113. Wang, P. What Do You Mean by «AI»?.. Proceedings of the First AGI Conference. 2008. P. 362- 373. URL: https://www.researchgate.net/publication/262357941_What_Do_You_Mean_by_AI (дата звернення: 28.07.2024).
114. Wang, Pei, Chen K, Yang S, Lin P. Relationship of sleep quality, smartphone dependence, and health-related behaviors in female junior college students. PLoS One. 2019. 14(4). URL: https://www.researchgate.net/publication/332207042_Relationship_of_sleep_quality_smartphone_dependence_and_health-related_behaviors_in_female_junior_college_students (дата звернення: 17.07.2024).
115. Wang, Pei. On Defining Artificial Intelligence. Journal of Artificial General Intelligence. 2019. Т.10. Вип. 2. С1-37. URL: <https://doi.org/10.2478/jagi-2019-0002> (дата звернення: 17.07.2024).
116. What Is A Deepfake? URL: <https://www.fortinet.com/resources/cyberglossary/deepfake> (дата звернення: 07.10.2024).

117. When ethics of Artificial Intelligence. URL: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics> (дата звернення: 02.08.2024).

118. When Global Partnership on Artificial Intelligence. URL: <https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html> (дата звернення: 02.08.2024).

119. When shaping the future of artificial intelligence: Reflections from the AI Governance Alliance. URL: <https://www.weforum.org/agenda/2024/06/ai-governance-alliance-future-of-ai/> (дата звернення: 02.08.2024).

120. When UN Secretary-General Launches Advisory Board to Support AI Governance. URL: <https://sdg.iisd.org/news/un-secretary-general-launches-advisory-board-to-support-ai-governance/> (дата звернення: 02.08.2024).

121. Whitney L. Are computers already smarter than humans. Time. September 29. 2017. URL: <https://time.com/4960778/computers-smarter-than-humans> (дата звернення: 24.09.2024).

122. Wilkinson, Jaci. "Artificial Unintelligence." Journal of Web Librarianship 13(2). 2019. 199. URL: <https://www.tandfonline.com/doi/full/10.1080/19322909.2019.1586626> (дата звернення: 24.09.2024).

123. Wright N. How artificial intelligence will reshape the global order. Foreign Affairs. 2018. URL: <https://www.foreignaffairs.com/articles/world/2018-07-10/how-artificial-intelligence-will-reshape-global-order> (дата звернення: 10.09.2024).

124. Yantaç C. E. Artificial intelligence, society and democracy. Turkish Policy Quarterly. 2021. URL: <http://turkishpolicy.com/article/1097/artificial-intelligence-society-and-democracy> (дата звернення: 10.09.2024).

125. Živković Katarina, Dušan Đurić. The AI Act – EU's First Artificial Intelligence Regulation. KINSELLAR. January 2024. URL:

<https://www.kinstellar.com/news-and-insights/detail/2577/the-ai-act-eus-first-artificial-intelligence-regulation> (дата звернення 19.11.2024).

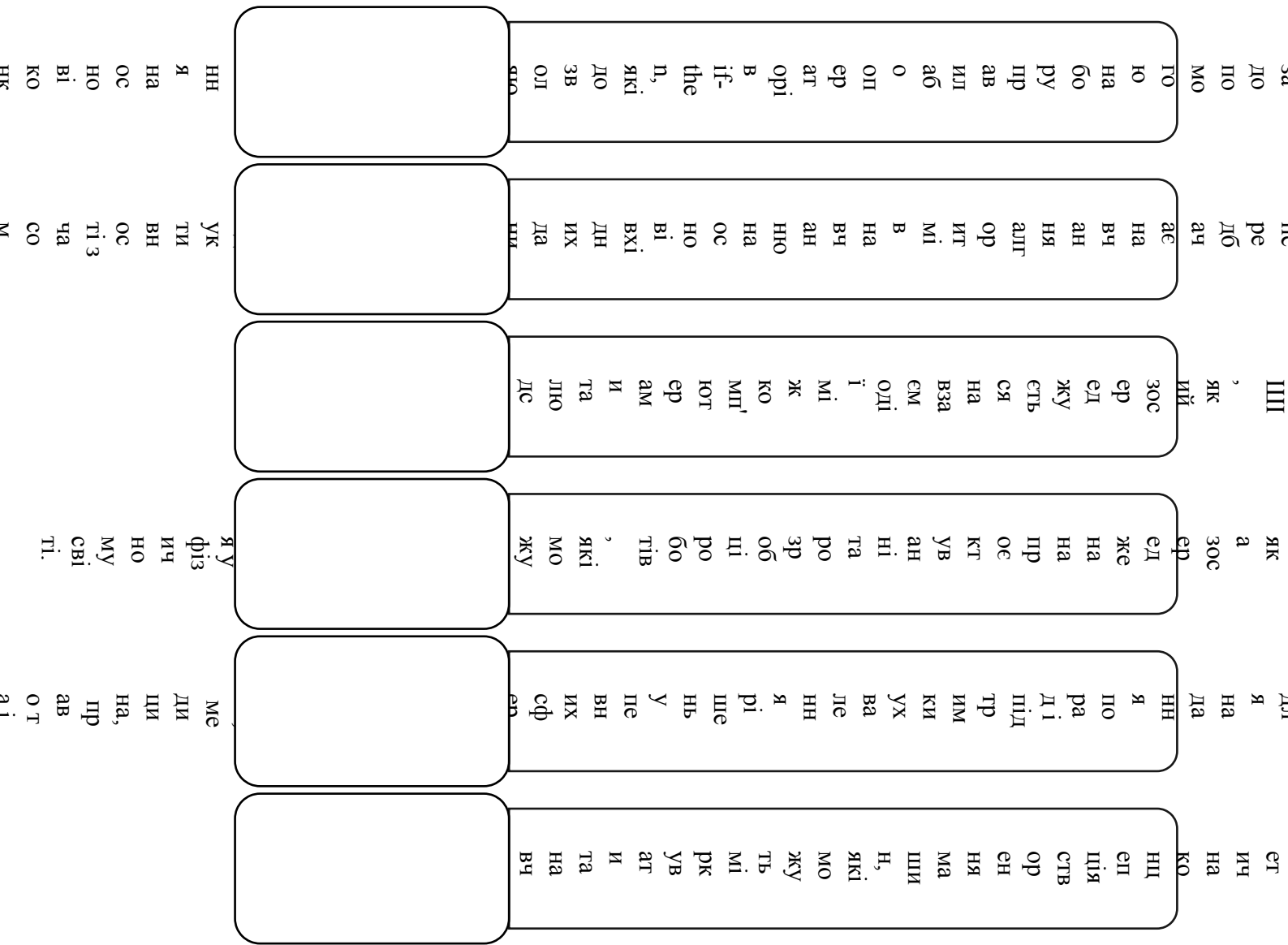
ДОДАТКИ

Додаток 1

Основні думки про мету заснування ІІІ [38]



Типи Штучного інтелекту [6]



Додаток 3

Стратегії зменшення впливу ШІ на права людини, демократію та
верховенство права



Список журналів та статей присвячених дослідженню Штучного
інтелекту [2]

1. Штучний інтелект, видавництво Elsevier.
2. Журнал досліджень машинного навчання, видавництво Мікротом.
3. Автономні агенти та мультиагентні системи, видавництво Спрингер.
4. Транзакції IEEE на нейронних мережах і системах навчання.
5. Журнал досліджень Штучного інтелекту, видавництво AAAI.
6. Машинне навчання, видавництво Спрингер.
7. Обчислювальний інтелект, видавництво Вайлі-Блеквелла.
8. Експертні системи, видавництво Вайлі-Блеквелла.
9. Міжнародний журнал інтелектуальних систем, видавництво Wiley.
10. Нейрокомп'ютинг, видавництво Elsevier.
11. Журнал експериментального та теоретичного Штучного інтелекту, видавництво Тейлор і Френсіс.
12. Журнал IEEE computational intelligence, видавництво IEEE.
13. Огляд Штучного інтелекту, видавництво Спрингер.
14. Автономні роботи, видавництво Спрингер.
15. Міжнародний журнал машинного навчання та кібернетики, видавництво Спрингер.
16. Операції ACM щодо інтелектуальних систем і технологій, видавництво Асоціація обчислювальної техніки.
17. Прогрес у Штучному інтелекті, видавництво Спрингер.
18. Ройовий інтелект, видавництво Спрингер.

SUMMARY

The term "Artificial Intelligence" has both broad and narrow meanings. In a narrow sense, it refers to information created by automated systems. In general, AI signifies the ability of a system to perform human functions by analyzing factors and making decisions based on experience. John McCarthy first defined AI in 1956 as the creation of agents that adapt to new situations and achieve goals unattainable by traditional systems. The High-Level Expert Group on Artificial Intelligence of the European Commission defines AI as a system operating in the physical or digital world, analyzing data to achieve specific objectives. According to the Joint Research Centre, AI is a system that recognizes its environment and data to select the best course of action.

The risks of Artificial Intelligence include several major threats. Firstly, malicious actors can use generative AI to scale attacks, improve ransomware methods, and execute phishing schemes. They can manipulate autonomous systems such as driverless cars or medical equipment, endangering physical safety. Secondly, automated malware can be developed using AI, such as ChatGPT, potentially displacing software developers. Privacy risks include leaks of personal information due to errors or attacks on AI systems. There is also the threat of theft and manipulation of AI models, which could intensify malicious activities. Another issue is data manipulation, which may lead to undesirable or harmful outcomes. AI-generated personifications can be exploited for virtual fraud and fake records.

Among the advantages of AI is its proactive ability to prevent cyberattacks by identifying and neutralizing threats before they cause harm. It enhances incident response by quickly analyzing data and automatically mitigating threats in real time. AI also improves threat detection accuracy, reduces human error, and frees up time for more strategic work. It efficiently processes large volumes of data, identifying anomalies and network issues, and streamlines policy development and procedures. AI also enhances cybersecurity training by creating realistic attack simulations, provides scalability to accommodate organizational growth, and adapts to changes in

cyber risks. Automated testing and vulnerability monitoring ensure continuous security checks. AI is also used for risk prediction, user authentication, and fraud detection.

Artificial Intelligence holds immense potential to improve society but also presents significant challenges to human rights, democracy, and the rule of law. Its impact is both positive and negative.

AI can be used to predict crimes or assess the likelihood of reoffending, affecting the right to liberty and security. It is widely employed for mass surveillance, facial recognition, and biometric data processing, threatening the right to privacy (Article 8 of the European Convention on Human Rights). AI may perpetuate or amplify discriminatory biases, as evidenced by Amazon's recruitment system, which favored men due to biased data. AI in recommendation systems, news aggregators, and search engines can limit access to information or deliver personalized yet biased or falsified information. This affects democratic discourse, fostering "information echo chambers" and manipulating public opinion.

Personalized advertising and microtargeting can manipulate voters by delivering tailored messages to specific groups, undermining the fairness and transparency of the electoral process. AI can also be used to create deepfakes and disinformation, complicating the identification of truthful information and eroding trust in media and democratic processes.

Decision-making automation in critical sectors such as finance, energy, and transportation can create new systemic risks and threats to societal systems. AI may worsen working conditions, infringe on the right to decent work and equal treatment, especially when used to monitor employees or assess their performance.

Addressing these issues requires regulatory measures. The European Union, the United States, and the United Kingdom have introduced various initiatives and legal acts to regulate AI implementation, minimize risks, and enhance transparency. The EU, in particular, has adopted the Artificial Intelligence Act, establishing harmonized rules for risk management and ethical technology use. In the U.S.,

recommendations from the National Institute of Standards and Technology (NIST) focus on risk management and responsible AI deployment. In the U.K., existing legal norms are being adapted to account for digital technology specifics, including AI's impact on workers' rights and cybersecurity.

A key component of these initiatives is integrating principles of ethics, transparency, and human rights protection, enabling nations to effectively address challenges arising from automation. Such approaches ensure the safe use of AI in critical sectors like finance, energy, and transportation. At the same time, these initiatives promote international cooperation to establish global standards for AI regulation, preventing fragmented approaches and ensuring balanced integration of innovations across countries.

In Ukraine, the 2020 Concept for the Development of Artificial Intelligence outlines the goals, principles, and objectives for AI technologies. The National Association of Lawyers of Ukraine has formed a Working Group to analyze legal issues related to AI, including personal data protection. Amendments to copyright law concerning AI-created works were introduced in December 2022, establishing sui generis rights. In 2024, the Ministry of Digital Transformation developed a White Paper describing AI regulatory approaches. The document is designed to help companies comply with new requirements and ensure product safety. Business guidelines, codes of conduct, and a regulatory sandbox have been developed.

The UN General Assembly adopted a resolution on March 11, 2024, regulating the development, production, and use of Artificial Intelligence. The European Parliament also adopted the AI Act in March 2024.

The Council of Europe's Ad Hoc Committee on Artificial Intelligence (CAI) has drafted the Convention on Artificial Intelligence (AI Convention), establishing general principles, including human dignity, transparency, accountability, and privacy protection. The Convention emphasizes safe innovation, reliability, and equal opportunities. Chapter IV addresses remedies and procedural safeguards, while Chapter VI focuses on promoting digital literacy and skills. The Convention aims to

foster innovation while respecting human rights and the rule of law. The EU Commission participated in the negotiations, and the Convention was adopted in March 2024.