Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна
Навчально-науковий інститут комп'ютерних наук та штучного інтелекту

Кафедра кібербезпеки інформаційних систем, мереж і технологій

До захисту допущено
Кафедрою КІСМіТ протокол №____ від «___» грудня 2025 р.

завідувач кафедри _____ Марина ЄСІНА
(підпис) *(ім'я, прізвище)*
«___» грудня 2025 р.

Кваліфікаційна робота

здобувача другого (магістерського) рівня вищої освіти
«Дослідження та аналіз міжнародних стандартів та регуляторних вимог щодо безпеки штучного інтелекту, розробка пропозицій та рекомендацій для України»

Спеціальність (спеціалізація) 125 «Кібербезпека та захист інформації»

Освітня програма «Безпека інформаційних і комунікаційних систем»

Виконавець _____ Єлизавета ЛОГАЧОВА
(підпис) *(ім'я, прізвище)*

Науковий керівник _____ Марина ЄСІНА
(підпис) *(ім'я, прізвище)*

Харків – 2025

РЕФЕРАТ

У роботі наведено: 9 рисунків, 35 таблиць, 32 джерела, 2 додатки. Обсяг роботи становить 63 сторінки.

Метою роботи є дослідження та аналіз міжнародних стандартів та регуляторних вимог щодо безпеки та використання штучного інтелекту, а також розробка пропозиції та рекомендацій для формування ефективної моделі безпеки штучного інтелекту в Україні.

Об’єкт дослідження – дослідження та аналіз міжнародних стандартів та регуляторних вимог у сфері безпеки штучного інтелекту.

Предмет дослідження – особливості застосування цих стандартів і вимог, механізми їх імплементації, а також можливості адаптації та використання в українському правовому та технологічному середовищі.

Методи дослідження – аналіз та порівняння відомих міжнародних регуляторних актів та методик з регулювання безпеки технологій штучного інтелекту; удосконалена методика оцінювання для порівняння різних підходів за обраними умовними та безумовними критеріями.

У роботі досліджено: міжнародні регуляторні акти та методики щодо регулювання технологій штучного інтелекту; українське нормативно-правове поле у напрямку безпечного використання штучного інтелекту; перспективи розвитку регуляторних вимог в Україні та перспективність імплементації міжнародного досвіду в Україні.

Результати роботи можуть бути використані у подальшому дослідженні даної області, у розробці регуляторних норм для України. А також у різних наукових виданнях.

Ключові слова: ШТУЧНИЙ ІНТЕЛЕКТ, БЕЗПЕКА, КІБЕРБЕЗПЕКА, СТАНДАРТИ, РЕГУЛЯТОРНІ ВИМОГИ, МАШИННЕ НАВЧАННЯ.

ABSTRACT

The paper contains: 9 figure, 35 tables, 32 sources, 2 application. The volume of the work is 63 pages.

The purpose of the work is to study and analyse international standards and regulatory requirements for the security and use of artificial intelligence, as well as to develop proposals and recommendations for the formation of an effective artificial intelligence security model in Ukraine.

The object of the study is the research and analysis of international standards and regulatory requirements in the field of artificial intelligence security.

The subject of the study is the specifics of the application of these standards and requirements, the mechanisms for their implementation, as well as the possibilities for their adaptation and use in the Ukrainian legal and technological environment.

The research methods are analysis and comparison of well-known international regulatory acts and methodologies for regulating the security of artificial intelligence technologies; an improved assessment methodology for comparing different approaches based on selected conditional and unconditional criteria.

The work examines: international regulatory acts and methodologies for regulating artificial intelligence technologies; the Ukrainian regulatory and legal framework for the safe use of artificial intelligence; prospects for the development of regulatory requirements in Ukraine and the prospects for implementing international experience in Ukraine.

The results of the work can be used in further research in this area, in the development of regulatory norms for Ukraine, and in various scientific publications.

Keywords: ARTIFICIAL INTELLIGENCE, SECURITY, CYBERSECURITY, STANDARDS, REGULATORY REQUIREMENTS, MACHINE LEARNING.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ 6
ВСТУП 7
1 ВПЛИВ ШТУЧНОГО ІНТЕЛЕКТУ НА РОЗВИТОК КІБЕРБЕЗПЕКИ 9
1.1 Загальні відомості про розвиток штучного інтелекту 9
1.1.1 Типи штучного інтелекту 10
1.1.2 Теорія роботи штучного інтелекту 12
1.2 Ризики та переваги використання ШІ 13
1.3 Огляд стану кібербезпеки в епоху штучного інтелекту 14
2 ПІДХОДИ ДО РЕГУЛЮВАННЯ ШІ У СВІТІ 17
2.1 Огляд міжнародних стандартів 17
2.2 Порівняльний аналіз підходів 26
2.3 Теоретичний огляд можливості інтеграції міжнародних підходів в Україні 27
3 РЕКОМЕНДАЦІЇ ТА СТВОРЕННЯ МОДЕЛІ БЕЗПЕКИ ШІ ДЛЯ УКРАЇНИ .. 30
3.1 Огляд поточної ситуації в українському правовому полі 30
3.2 Модель безпеки ШІ в Україні 31
3.3 Рекомендації та застереження для України 35
3.4 Оцінка адекватності застосування міжнародних стандартів та регуляторних актів для України 37
3.4.1 Вибір методу оцінювання та програмно-апаратного забезпечення 37
3.4.2 Опис послідовності дій для оцінювання 38
3.4.3 Опис та результати дослідження за сукупністю безумовних критеріїв .. 38
3.4.4 Опис та результати дослідження за сукупністю умовних критеріїв 40
3.4.5 Оцінювання адекватності застосування міжнародних методів методом визначення вагових коефіцієнтів за допомогою шкали Фішберна 45

3.4.6 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі методу ранжування 49

3.4.7 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі методу приписування балів..... 54

3.4.8 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі числового способу..... 58

3.4.9 Висновки за результатами оцінювання..... 59

ВИСНОВКИ..... 62

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ..... 64

ДОДАТОК А..... 68

ДОДАТОК Б 71

ПЕРЕЛІК СКОРОЧЕНЬ

ЄС – Європейський Союз

ОАЕ – Об’єднані Арабські Емірати

США – Сполучені Штати Америки

ШІ (AI) – штучний інтелект (artificial intelligence)

AIDA – Artificial Intelligence and Data Act, Закон про штучний інтелект та дані

DIFC – Dubai International Financial Centre, Дубайський міжнародний фінансовий центр

FCA – Financial Conduct Authority, Управління з фінансового контролю та нагляду

GPT-3 – Generative Pre-trained Transformer 3, Генеративний попередньо навчений перетворювач 3

GPT-4 – Generative Pre-trained Transformer 4, Генеративний попередньо навчений перетворювач 4

IBM – International Business Machines Corporation, Міжнародна корпорація бізнес-машин

IEC – International Electrotechnical Commission, Міжнародна електротехнічна комісія

ISO – International Organization for Standardization, Міжнародна організація зі стандартизації

NIST – National Institute of Standards and Technology, Національний інститут стандартів і технологій

RMF – Risk Management Framework, Структура управління ризиками

ВСТУП

За останні роки технології штучного інтелекту (ШІ) швидко розвинулись та набувають все ширшого застосування у різних сферах життя. Ці технології внесли великі зміни не тільки у процеси сервісу, освіти, побуту, а й не обійшли стороною інформаційну безпеку та безпеку загалом.

Технології штучного інтелекту, як і будь-які сучасні прориви створені для полегшення життя, швидко набули популярності серед мільйонів користувачів. Але там, де звичайний користувач бачить чергову розвагу, зловмисник бачить нову можливість.

Сучасні технології на основі розумних машин потребують великої уваги з точки зору безпеки. Адже однією з найскладніших задач є збереження балансу між можливістю подальшого розвитку технологій та збереженням прав і свобод звичайних користувачів та громадян.

Проте технології штучного інтелекту допомагають і у прогресі систем та алгоритмів безпеки. За даними зі звіту IBM [1] про вартість витоку даних за 2025 рік вартість витоку даних знизилась. І вперше за п’ять років показала нижчий показник, що за підрахунками становив 4,44 млн доларів. Разом з тим, 13% компаній повідомили про витоки даних через їх моделі штучного інтелекту. Використання ж несанкціонованих у компаніях моделей, тобто тіньового штучного інтелекту, збільшувало збитки приблизно на 670 тис. доларів [1].

Щоб не відмовлятись від прогресу, багато країн уже запровадили або почали запроваджувати регуляторні вимоги щодо використання штучного інтелекту у різних галузях. У деяких країнах було створено спеціальні органи для врегулювання питань пов’язаних з цими технологіями.

Для України, яка інтегрується у світовий технологічний та правовий простір, надзвичайно важливо проаналізувати ці практики та виробити власну модель регулювання безпеки штучного інтелекту. Це дозволить не лише адаптувати інновації без шкоди для суспільства, але й створити передумови для сталого розвитку цифрової економіки та захисту національних інтересів.

Тож актуальність даної роботи полягає у необхідності всебічного дослідження міжнародних стандартів і регуляторних вимог у сфері безпеки штучного інтелекту та їх адаптації до українського контексту при розробці власних механізмів регуляції. Відсутність належного регулювання може призвести до зростання кіберзагроз, порушень прав людини, неконтрольованого використання технологій і, навіть, ризику втрати довіри суспільства до цифрових рішень.

Об’єкт дослідження – дослідження та аналіз міжнародних стандартів та регуляторних вимог у сфері безпеки штучного інтелекту.

Предмет дослідження – особливості застосування цих стандартів і вимог, механізми їх імплементації, а також можливості адаптації та використання в українському правовому та технологічному середовищі.

Метою роботи є дослідження та аналіз міжнародних стандартів та регуляторних вимог щодо безпеки та використання штучного інтелекту, а також розробка пропозиції та рекомендацій для формування ефективної моделі безпеки штучного інтелекту в Україні.

Отримані результати дослідження будуть корисними для подальшого розвитку регулювання безпечного використання технологій штучного інтелекту. А також дозволять пришвидшити інтеграцію міжнародних стандартів у галузі безпеки в Україні. Результати також будуть корисними для звичайних користувачів технологій, пов’язаних зі штучним інтелектом, покращать їх обізнаність у цій галузі та допоможуть більш відповідально використовувати ці технології. При виконанні роботи було поставлено наступні задачі: проаналізувати ризики використання ШІ та пов’язані з цим загрози для кібербезпеки і захисту персональних даних; дослідити міжнародні стандарти та регуляторні підходи до безпеки ШІ; здійснити порівняльний аналіз практик різних країн у сфері регулювання ШІ; оцінити можливості імплементації міжнародного досвіду в Україні; вивчити сучасний стан правового регулювання використання ШІ в Україні; розробити рекомендації та пропозиції щодо формування моделі безпеки штучного інтелекту для України.

1 ВПЛИВ ШТУЧНОГО ІНТЕЛЕКТУ НА РОЗВИТОК КІБЕРБЕЗПЕКИ

1.1 Загальні відомості про розвиток штучного інтелекту

Поняття штучного інтелекту з’являлося в науковій фантастиці протягом багатьох років, коли автори досліджували різні сценарії розвитку інтелектуальних машин і технологій. Деякі закінчуються катастрофою, інші надихають розробників і вчених, зрештою перетворюючи вигадку на реальність. Десять років тому ми й уявити не могли, які можливості ШІ пропонує нам сьогодні. За своєю суттю ШІ відноситься до комп’ютерних систем, які можуть виконувати завдання, які зазвичай вимагають людського інтелекту, наприклад, міркування, навчання, сприйняття та розуміння мови. Ці системи аналізують величезні масиви даних, розпізнають закономірності та приймають рішення з безпрецедентною швидкістю та точністю. Застосування штучного інтелекту величезне та різноманітне – від побутової моделі, яка передбачає ваші потреби, до розробки ліків за допомогою штучного інтелекту, що прискорює медичні прориви.

Багато хто приписує початковий поштовх для розвитку цієї галузі різницевим та аналітичним механізмам Чарльза Беббіджа, які він винайшов у 19 столітті. На відміну від сучасних калькуляторів, різницевий механізм був розроблений не для виконання звичайних повсякденних арифметичних операцій, а для обчислення серії числових значень і автоматичного друку результатів, що є важливою віхою в історії обчислювальної техніки [2].

Беббідж використав принцип скінченних різниць, який у той час використовували людські комп’ютери для складання математичних таблиць. Він включав лише додавання та віднімання, які було набагато легше механізувати, ніж множення. Після того, як цей проєкт був раптово залишений, Беббідж задумав більш амбітну та технічно складнішу машину. Аналітичний механізм був розроблений для виконання будь-якого набору попередніх обчислень і мав навіть більші аналітичні можливості, ніж оригінальний різницевий механізм. Цей

механізм вважається першою повністю автоматичною обчислювальною машиною [2].

Британського зломщика кодів Алана Тюрінга, який був ключовою фігурою в розвідувальному арсеналі союзників під час Другої світової війни, також можна вважати батьком сучасних ітерацій ШІ. У 1950 році він запропонував тест Тюрінга «Гра в імітацію», призначений для оцінки здатності машини демонструвати розумну поведінку, яка не відрізняється від поведінки людини [3].

З цього моменту розвиток технології штучного інтелекту почав експоненційно прискорюватися і його очолили такі впливові особи, як Джон Маккарті, Йозеф Вайценбаум, Герберт Саймон, Едвард Файгенбаум та Джошуа Ледерберг, Джеймс Лайтхілл, та багато інших [3].

У 1997 році комп’ютер IBM Deep Blue переміг чемпіона світу Гаррі Каспарова у гучному шаховому матчі, продемонструвавши величезний потенціал систем штучного інтелекту. Того ж року на платформі Windows було реалізовано програмне забезпечення для розпізнавання мови, розроблене Dragon Systems. У 2011 році комп’ютер IBM Watson Deep QA переміг двох беззаперечних чемпіонів у вікторині, продемонструвавши здатність систем штучного інтелекту розуміти природну мову [3].

У 2017 році прогрес у глибокому навчанні та збільшенні обчислювальної потужності сприяли швидкому розвитку комп’ютерного зору, обробки природної мови, робототехніки та автономних систем. Через шість років, у 2023 році, світ був вражений появою великомасштабних мовних моделей, таких як GPT-3 та її наступників, які продемонстрували потенціал систем штучного інтелекту генерувати текст, схожий на людину, відповідати на запитання та допомагати з широким спектром завдань [3].

1.1.1 Типи штучного інтелекту

Ми звикли узагальнено використовувати назву «штучний інтелект» для загального окреслення поняття розумних машин, проте насправді ж штучний інтелект поділяється на чотири типи в залежності від сфери його застосування та можливостей.

Реактивні машини – це системи штучного інтелекту, що працюють суворо за заздалегідь визначеними правилами, але не мають можливості навчатися на нових даних або досвіді [4].

Яскравим прикладом є чат-боти, які генерують відповіді на основі запрограмованих алгоритмів, а не адаптуються до розмов [4]. Незважаючи на те, що вони відмінно справляються з конкретними завданнями, вони не можуть розвиватися далі свого початкового програмування.

Наступними є машини з обмеженою пам’яттю. На відміну від реактивних машин, системи штучного інтелекту з обмеженою пам’яттю можуть навчатися на історичних даних, що дозволяє їм приймати обґрунтовані рішення на основі минулого досвіду [4].

Ці типи штучного інтелекту спостерігаються в безпілотних автомобілях, які використовують датчики та алгоритми машинного навчання для аналізу моделей дорожнього руху та безпечної навігації в динамічному середовищі [4]. Подібним чином програми обробки природної мови використовують історичні дані для покращення розуміння та інтерпретації мови з часом.

Машини з теорією розуму поки є більш теоретичним типом, проте все ж ймовірним. Цей вид штучного інтелекту зможе розуміти людські емоції, наміри та соціальні ознаки [4]. Машина з теорією розуму могла б потім використовувати цю інформацію, щоб передбачати людські дії та вступати в інтуїтивну, емпатійну взаємодію [4]. Якщо це реалізувати, цей штучний інтелект міг би революціонізувати людино-комп’ютерну та соціальну робототехніку, створивши системи, які справді розуміють нас.

І останнім типом є – самосвідомий ШІ, що наразі є виключно теоретичний. Це є найбільш суперечливою концепцією, самосвідомий ШІ стосується машин із людською свідомістю, які усвідомлюють власне існування та здатні сприймати емоції інших [4].

Ці чотири типи ШІ підкреслюють широкий спектр інтелекту в штучних системах. У міру розвитку технологій штучного інтелекту вивчення можливостей і обмежень кожного типу допоможе краще корегувати використання даної

технології і допускати якомога менше злочинної діяльності пов’язаної з використанням штучного інтелекту.

1.1.2 Теорія роботи штучного інтелекту

За своєю суттю ШІ обробляє величезні обсяги даних, виявляючи закономірності та роблячи прогнози з надзвичайною точністю. Це досягається завдяки використанню великих наборів даних і інтелектуальних алгоритмів ШІ – структурованих наборів правил, які дозволяють програмному забезпеченню навчатися на шаблонах у даних. Рушійною силою є нейронні мережі – складні системи взаємопов’язаних вузлів, які передають інформацію через кілька рівнів, щоб знаходити зв’язки та витягувати значення з даних.

У основі штучного інтелекту лежать декілька важливих концепцій, їх можна виділити наступним чином. Машинне навчання, яке дозволяє системам аналізувати дані, розпізнавати закономірності та приймати рішення без явного програмування. А також механізм логічного висновку, що виступає в якості мозку системи міркування ШІ. Як правило у розумних машинах використовуються конкретні заздалегідь передбачувані типи міркувань, що врешті-решт визначають фінальний результат їх роботи [5].

ШІ використовує різні типи міркувань, їх застосування часто залежить від типу даних та очікуваного результату. Можна виділити наступні види міркувань [5]:

a) Абдуктивне міркування – полягає у виборі найімовірнішого висновку на основі наявних даних.
b) Агентне міркування – такий тип мислення дозволяє вирішувати завдання на основі попередніх прикладів, тобто діяти за попереднім алгоритмом.
c) Аналогічне міркування – у цьому випадку відбувається пошук аналогій з попереднього досвіду та наданих нових даних.
d) Здорове міркування – таке мислення покладається на загальновідому інформацію та повсякденні дані, наприклад, природню мову.

e) Дедуктивне міркування – дане мислення імітує експертне мислення. Тобто висуваючи припущення, що, якщо твердження є істинним, значить і висновки є такими ж.
f) Нечітке міркування – враховує ступені істини, більш широкі, ніж поняття «істина» чи «хиба».
g) Індуктивне міркування – використовує конкретні спостереження для отримання ширшого узагальнення для своїх висновків.
h) Нейросимволічне міркування – апелює символами, а не числами.
i) Імовірнісне міркування – використовує ймовірності для прийняття рішень, оцінюючи їх.
j) Просторове міркування – націлене на сприйняття поверхонь та вимірів і пояснення роботи у них та з ними.
k) Часове міркування – дозволяє працювати з конкретними періодами, розуміти часові послідовності та наслідковості.

Та все ж усі типи мислення упираються у кілька проблем, що зрештою можуть виникати при роботі з ШІ, а саме: упередженість, обчислювальні втрати та інтерпретовність.

1.2 Ризики та переваги використання ШІ

Штучний інтелект пропонує широкий спектр можливостей для вдосконалення процесів, оптимізації якості обслуговування та підвищення ефективності в численних галузях. Його впровадження охоплює різні аспекти, включаючи медицину, фінансовий сектор, виробничу індустрію та сферу розваг. Утім, поряд із суттєвими перевагами виникають і нові виклики, пов'язані з питаннями безпеки, етичними нормами та захистом конфіденційної інформації.

Серед ключових переваг, які забезпечує впровадження штучного інтелекту, можна виділити такі аспекти, як: аналіз великих масивів даних, автоматизація рутинних завдань, оптимізація прийняття рішень, підвищення рівня обслуговування, запровадження інновацій у різноманітних сферах, оперативніше реагування на кіберзагрози, вдосконалена аналітика ризиків та багато інших можливостей [7].

Втім, використання ШІ також супроводжується значним спектром викликів і ризиків. До них належать: порушення конфіденційності, алгоритмічна упередженість, неетичне застосування технологій, потенційні втрати робочих місць, автоматизація кіберзагроз, відсутність довіри та складнощі з контролем, використання штучного інтелекту для соціальних маніпуляцій, атаки на власні алгоритми ШІ та чимало інших небезпек [7].

Закон про штучний інтелект [7, 8], ухвалений Європейським парламентом, класифікує ризики, пов'язані з ШІ, на високі й неприйнятні. Системи ШІ, віднесені до категорії неприйнятного ризику, розглядаються як загроза для людства й підлягатимуть забороні. До них належать [7]:

• когнітивне поведінкове маніпулювання людьми або окремими вразливими групами;

• соціальна оцінка: класифікація людей на основі поведінки, соціально-економічного статусу чи особистих характеристик;

• біометрична ідентифікація та категоризація людей;

• системи біометричної ідентифікації в режимі реального часу та віддалені, наприклад, розпізнавання обличчя.

Окремо важливо підкреслити, що в деяких випадках можуть бути дозволені винятки для потреб правоохоронної діяльності [7].

Системи штучного інтелекту, які створюють загрозу безпеці або порушують фундаментальні права, класифікуються як високоризикові та поділяються на дві основні категорії. Перша охоплює штучний інтелект, інтегрований у продукти, що регулюються законодавством Європейського Союзу щодо безпеки. До таких продуктів належать, зокрема, іграшки, авіаційна техніка, автомобілі, медичні пристрої та ліфти. Друга категорія включає системи штучного інтелекту, які задіяні у визначених сферах діяльності та повинні бути зареєстровані у відповідній базі даних ЄС [6].

1.3 Огляд стану кібербезпеки в епоху штучного інтелекту

Штучний інтелект кардинально змінює цифровий світ, впроваджуючи автоматизацію процесів, аналіз великих обсягів даних та прогнозування

потенційних загроз. Якщо раніше його застосовували для виконання простих і повторюваних завдань, то сьогодні він активно використовується для медичної діагностики, оптимізації бізнес-процесів, освітніх потреб та багато іншого. Технологія розвивається з вражаючою швидкістю й охоплює майже всі аспекти повсякденного життя, значно спрощуючи його для користувачів.

За даними IBM, у 2023 році середня вартість витоку даних перевищувала 4,45 мільйона доларів, а штучний інтелект поступово ставав ключовим інструментом як для захисту, так і для атак. У 2024 році, згідно зі звітом IBM Cost of a Data Breach Report, ця сума зросла до 4,88 мільйона доларів США, що на 10% більше порівняно з попереднім роком [6]. Водночас завдяки штучному інтелекту та автоматизації компанії змогли скоротити витрати на усунення наслідків витоків даних у середньому на 2,2 мільйона доларів. ШІ ефективно вирішує низку критичних завдань у сфері кібербезпеки, таких як виявлення загроз у режимі реального часу, автоматизоване реагування на інциденти, прогнозування атак та інших викликів [6].

З розвитком штучного інтелекту зростають й нові загрози, особливо у сфері кібербезпеки. Зловмисники все частіше звертаються до можливостей ШІ для створення контенту типу deepfake, автоматизації фішингових атак і проведення соціальної інженерії на новому рівні. Крім того, штучний інтелект може використовуватися для автоматизації шкідливого програмного забезпечення. Одним із яскравих прикладів є експеримент, проведений фахівцями кібербезпеки компанії Hyas. У ході дослідження вони створили умовний вірус під назвою BlackMamba, який здатний динамічно адаптувати свою поведінку за допомогою ChatGPT [6].

Цей код генерував частини свого функціоналу в реальному часі, демонструючи високу здатність до адаптації у різних середовищах і залишаючись непомітним для більшості популярних антивірусних систем під час тестування [6].

У лютому 2024 року в Гонконзі стався один із наймасштабніших інцидентів шахрайства з використанням штучного інтелекту. Співробітнику міжнародної

компанії доручили перевести 25 мільйонів доларів США від імені начебто директора з Великої Британії. Початково підозра щодо фішингу виникла після отримання сумнівного листа із запрошенням на відеоконференцію. Однак поява кількох «знайомих колег» під час дзвінка послабила його настороженість [6].

У результаті через серію транзакцій, задіявши п'ять різних банківських рахунків, він перевів загалом понад 200 мільйонів гонконзьких доларів (приблизно 25 мільйонів доларів США). Згодом стало відомо, що всі інші учасники відеоконференції були надзвичайно реалістичними цифровими двійниками, створеними зловмисниками за допомогою технологій штучного інтелекту та deepfake [6].

Це створює складний виклик для фахівців з інформаційної безпеки, оскільки новітня технологія не лише значно зміцнює безпеку, але й може бути використана кіберзлочинцями для здійснення атак. Оптимальним підходом стає запровадження відповідних стандартів і регуляторних норм для забезпечення безпечного застосування штучного інтелекту. Особливо актуальним це питання є для України, де в умовах війни стрімко зростає кількість фейків та кібератак [6].

2 ПІДХОДИ ДО РЕГУЛЮВАННЯ ШІ У СВІТІ

2.1 Огляд міжнародних стандартів

Існує широкий спектр підходів та поглядів на нормативно-правове регулювання штучного інтелекту. У деяких країнах до цієї технології ставляться з обережністю, зважаючи на можливі ризики, тоді як інші бачать у ній значний прогресивний стимул для розвитку цифрової трансформації.

Особливу увагу варто звернути на Європейський Союз та їх EU AI Act – перший офіційний та комплексний європейський документ, спрямований на регулювання сфери штучного інтелекту. Цей акт запроваджує чіткі правила використання відповідних технологій, базуючись на оцінці рівня потенційного ризику, який може супроводжувати різні системи ШІ. Залежно від ступеня цього ризику як постачальники технологій, так і користувачі зобов'язані дотримуватися встановлених вимог. Навіть для систем із незначним ризиком передбачено обов’язкове оцінювання, що має підтвердити їхню прозорість і відповідність етичним стандартам [6, 8].

До категорії недопустимого ризику відносяться програми штучного інтелекту, які в Європейському Союзі заборонені через загрозу основним правам і свободам людини. Сюди належать технології, що маніпулюють поведінкою людей, особливо уразливих груп, таких як діти. Як приклад, можна навести інтерактивні голосові іграшки, здатні спонукати до небезпечних дій. Крім того, до цього списку входять системи соціального рейтингування, які оцінюють людей за їхньою поведінкою чи статусом, біометрична ідентифікація в реальному часі у громадських місцях, а також класифікація осіб на основі біометричних характеристик [6, 8].

У певних ситуаціях, зокрема під час роботи правоохоронних органів, можливе обмежене використання високоризикових систем штучного інтелекту. Наприклад, дистанційна біометрична ідентифікація в реальному часі дозволяється лише у виняткових випадках і тільки після попереднього рішення суду.

Ідентифікація «постфактум», тобто із затримкою, може застосовуватися під час розслідування тяжких злочинів, але також виключно за наявності юридичного дозволу [6, 8].

Системи штучного інтелекту з високим рівнем ризику охоплюють сфери, що мають безпосередній вплив на безпеку та фундаментальні права громадян. Вони поділяються на дві основні категорії. Перша – це ШІ, інтегрований у продукти, які регулюються європейським законодавством щодо безпеки, наприклад, медичне обладнання, транспортні засоби, іграшки та ліфти [6, 8].

Друга – це системи, що функціонують у чутливих галузях, таких як управління критичною інфраструктурою, освіта, працевлаштування, доступ до державних послуг, діяльність правоохоронних органів, а також процеси міграції та правозастосування [6, 8].

Усі ці системи підлягають обов’язковому контролю як перед виходом на ринок, так і протягом усього періоду їх використання. Крім того, громадяни матимуть право подавати скарги на роботу таких ШІ-систем до національних наглядових органів. Це зміцнює принципи прозорості та підзвітності в умовах технологічного розвитку.

Одним із ключових принципів є прозорість. Так, наприклад, ChatGPT не належить до систем із високим рівнем ризику, проте повинен дотримуватися законодавства ЄС, зокрема вимог щодо авторського права. Тому весь контент, створений ChatGPT, має бути спеціально маркований – це допомагає уникати поширення фейкових зображень, відео та іншого контенту, який може використовуватися з шахрайською метою. Більш розвинена модель штучного інтелекту GPT-4 повинна проходити ретельну оцінку, а про будь-які серйозні інциденти необхідно повідомляти Європейську комісію [6, 8].

Перші норми, зокрема заборона на використання систем із неприйнятним рівнем ризику, набули чинності 2 лютого 2025 року. Інші положення, такі як кодекси практики та вимоги до прозорості універсальних моделей ШІ, почнуть діяти через дев’ять і дванадцять місяців відповідно [6].

Системи з високим рівнем ризику, що застосовуються у сфері охорони здоров’я, освіти, критичної інфраструктури або правоохоронній діяльності, матимуть більше часу для адаптації. Для них вимоги набудуть чинності через 36 місяців, що дасть змогу розробникам і користувачам належним чином підготуватися до впровадження нових стандартів [6].

NIST AI Risk Management Framework – це стратегічний документ, створений Національним інститутом стандартів і технологій США (NIST). Його мета – допомогти організаціям ефективно керувати ризиками, що виникають під час розроблення, впровадження та використання систем штучного інтелекту [6, 9].

Хоча застосування цього документа є добровільним, його широко рекомендують як інструмент, що сприяє створенню безпечних, надійних, прозорих і етично обґрунтованих рішень у сфері ШІ.

У межах документа виділяють три категорії потенційної шкоди: шкода людям, шкода організаціям та шкода екосистемі [6, 9].

Перша категорія – шкода людям – охоплює фізичні, психологічні, соціальні або економічні наслідки. Це може бути поширення фальшивих медичних порад, неправдивих новин чи контенту, що провокує насильство. Також штучний інтелект може використовуватися для дискримінації, упередженого оцінювання, незаконного спостереження або цензури, особливо коли рішення ухвалюються автоматично без належного людського контролю [6, 9].

Друга категорія – шкода організації. Вона включає ризики, пов’язані з витісненням творчих професій, недобросовісною конкуренцією через масове створення фальшивого контенту, дестабілізацією ринку праці, репутаційними втратами або іншими загрозами для бізнесу та безпеки [6, 9].

Третя категорія – шкода екосистемі. Сюди належать порушення у глобальних фінансових або логістичних системах, шкода довкіллю та природним ресурсам. Такий поділ допомагає зосередити увагу на попередженні ризиків і наслідків, що можуть виникнути у різних сферах впливу ШІ [6, 9].

Документ визначає сім основних характеристик, яким мають відповідати надійні системи штучного інтелекту.

Перші з них – валідність і надійність, тобто здатність системи виконувати поставлені завдання точно, стабільно та передбачувано навіть у змінних умовах. Такі системи повинні бути перевіреними та відповідати заявленим цілям. Безпечність означає, що система не повинна завдавати фізичної чи психологічної шкоди користувачам і має бути захищеною від зловмисного втручання [6, 9].

Наступна характеристика – неупередженість. ШІ повинен уникати дискримінації та приймати справедливі рішення незалежно від статі, віку, раси чи соціального статусу користувачів. Важливою є також прозорість, що передбачає відкритість щодо принципів роботи системи, джерел даних і можливих обмежень [6, 9].

Пояснюваність є ще однією ключовою вимогою – користувачі та розробники повинні мати змогу чітко пояснити, чому система ухвалила те чи інше рішення. Це особливо важливо у сферах, де від рішень ШІ залежать життя чи добробут людей – у медицині, фінансах або праві. Також важливою є захищеність конфіденційності, тобто дотримання правил збору, зберігання й обробки персональних даних відповідно до етичних і правових стандартів [6, 9].

Остання характеристика – стійкість. Вона означає здатність системи працювати навіть у разі збоїв, атак або непередбачуваних ситуацій [6, 9]. Надійні системи мають мати механізми самовідновлення та захисту від зовнішніх загроз.

NIST AI RMF ґрунтується на п’яти ключових функціях: Govern, Map, Measure, Manage та Improve. Разом вони формують єдину систему управління ризиками протягом усього життєвого циклу штучного інтелекту – від розробки до постійного вдосконалення [6, 9].

Перша функція – Govern – полягає у створенні внутрішньої структури управління: визначенні відповідальних осіб, прозорих процесів і політик, що регулюють використання ШІ [6, 9]. Це допомагає ухвалювати етично, юридично й соціально обґрунтовані рішення.

Друга функція – Map, що передбачає аналіз контексту, у якому працює система [6, 9]. Організації оцінюють можливі ризики, потреби користувачів і потенційний вплив ШІ на людей та довкілля.

Measure – це етап оцінювання ризиків і надійності системи. Використовуються кількісні та якісні показники, такі як точність, справедливість або стабільність [6, 9]. Регулярне вимірювання дозволяє вчасно виявляти слабкі місця й ухвалювати коригувальні рішення.

Функція Manage спрямована на практичне усунення або мінімізацію ризиків. Це може включати вдосконалення алгоритмів, обмеження доступу до певних функцій чи оновлення політик безпеки. Управління ризиками – це постійний процес, який триває протягом усього циклу використання ШІ [6, 9].

Остання функція – Improve, яка зосереджується на безперервному вдосконаленні системи. Організації повинні враховувати нові знання, технологічні зміни та оновлення нормативної бази, щоб підтримувати актуальність і безпечність своїх рішень [6, 9].

Додатково варто згадати міжнародний стандарт ISO/IEC 23894:2023. Це перший у світі документ, що надає рекомендації з управління ризиками, пов’язаними з використанням штучного інтелекту. Він адаптує загальні принципи менеджменту ризиків до особливостей ШІ – складності систем, етичних аспектів та впливу на права людини. Головна мета стандарту – мінімізувати можливу шкоду, підвищити довіру до штучного інтелекту й забезпечити його безпечне та передбачуване використання [6, 10].

У стандарті ISO/IEC 23894:2023 визначено структуру ефективного управління ризиками, пов’язаними зі штучним інтелектом, яка охоплює всі етапи життєвого циклу систем ШІ. Одним із перших і найважливіших етапів є ідентифікація ризиків, що передбачає глибоке розуміння практичного використання системи. На цьому етапі організації мають враховувати як заплановані сценарії застосування, так і можливі випадки неправильного або зловмисного використання [6, 10].

Важливо також проаналізувати якість і походження даних, на яких навчалася модель, принципи прийняття нею рішень, а також потенційний вплив її роботи на різні групи користувачів і суспільство загалом [6, 10].

Після визначення потенційних загроз стандарт передбачає проведення оцінки ризиків – як у кількісному, так і в якісному аспекті. Таке оцінювання має враховувати не лише ймовірність виникнення певної проблеми, а й масштаб можливих наслідків. Особливу увагу приділено аналізу каскадних ефектів, тобто ситуацій, коли один ризик може спричинити ланцюгову реакцію інших небажаних подій. Це дозволяє створити більш комплексне уявлення про потенційні загрози, які можуть виникати під час взаємодії системи ШІ з іншими технологіями або соціальними процесами [6, 10].

Наступний етап – обробка ризиків, тобто розроблення практичних рішень для їх зменшення чи нейтралізації. Це може включати зміни в архітектурі моделі, посилення технічних або організаційних заходів контролю, страхування ризиків або свідоме прийняття певного рівня залишкової небезпеки. Головне, щоб обрані методи відповідали природі ризику та узгоджувалися із загальною стратегією організації в питаннях етики, безпеки й відповідальності [6, 10].

Останнім елементом системи управління є безперервний моніторинг і перегляд ризиків. Оскільки системи штучного інтелекту постійно розвиваються та функціонують у змінному середовищі, стандарт ISO/IEC 23894 наголошує на необхідності регулярного оновлення оцінок ризиків. Це передбачає визначення ключових індикаторів ризику (KRI), перевірку ефективності вже впроваджених заходів і своєчасне коригування стратегій відповідно до технологічних або соціальних змін [6, 10].

Вже у 2025 році у США знову активно піднялось питання правильності регулювання розумних машин. «Розробка систем штучного інтелекту в Америці має бути вільною від ідеологічних упереджень чи штучно створених соціальних програм», – заявив Білий дім. «Завдяки правильній урядовій політиці Сполучені Штати можуть зміцнити свої позиції лідера у сфері штучного інтелекту та забезпечити світліше майбутнє для всіх американців» [11].

У 2023 році Джо Байден проводив свою політику щодо ШІ. Він підписав указ, що за яким мала регулюватись безпека використання штучного інтелекту в уряді. Трамп підписав виконавчий указ, спрямований на прискорення розвитку

штучного інтелекту, ліквідацію ідеологічних упереджень та впровадження сучасного плану дій у цій галузі, що фактично скасувало указ його попередника. В основному план спрямований на скасування попередніх політик, що обмежували розвиток штучного інтелекту [11].

Найбільш відкритою до впровадження технологій штучного інтелекту країною є Об'єднані Арабські Емірати (ОАЕ). Держава демонструє швидкий, стратегічний і водночас прагматичний підхід до розвитку та регулювання ШІ. Метою ОАЕ є не лише активне використання штучного інтелекту у своїй економіці, а й затвердження статусу глобального лідера в галузі технологічного управління [6, 12].

Ще у 2017 році ОАЕ стали першою країною у світі, яка призначила міністра штучного інтелекту. Цей крок став чітким сигналом політичної підтримки інноваційного розвитку. З того часу в державі реалізується Національна стратегія ШІ 2031, головна мета якої – зробити технології штучного інтелекту ключовим інструментом підвищення ефективності державного управління, системи охорони здоров'я, транспорту, освіти та інших суспільно важливих сфер [6, 12].

На відміну від багатьох інших країн, ОАЕ не запровадили єдиного детального законодавчого акту, подібного до Європейського AI Act. Замість цього держава застосовує галузевий підхід до регулювання, поєднуючи його з ініціативами публічно-приватного партнерства. Наприклад, у фінансовому секторі Центральний банк ОАЕ та інші регулятори вже впровадили спеціальні політики, що регулюють використання ШІ для запобігання шахрайству, зловживанням та упередженим рішенням у автоматизованих системах [6, 12].

Важливу роль у розвитку етичних стандартів ШІ відіграє Dubai International Financial Centre (DIFC) – спеціальна економічна зона з власною нормативною базою. У межах ініціативи DIFC AI and Data Protection Guidelines створено фреймворк для використання штучного інтелекту, який поєднує етичні норми, прозорість і підзвітність розробників [6, 12]. Таким чином, ОАЕ формують модель «гнучкого регулювання», що дозволяє адаптуватися до швидких технологічних змін без зайвої бюрократії.

Водночас відсутність єдиного національного закону про ШІ може призводити до фрагментації правового поля, особливо для міжнародних компаній, які працюють у різних еміратах або галузях. У майбутньому ОАЕ, ймовірно, розглянуть можливість прийняття більш цілісного законодавства, яке забезпечить єдині стандарти для всіх секторів та підвищить довіру бізнесу до етичного й безпечного впровадження технологій штучного інтелекту.

Албанія також має амбітні плани щодо впровадження штучного інтелекту навіть у державні структури. Politico зазначає, що Албанія обрала шлях розвитку штучного інтелекту, як інструмент до майбутнього вступу у ЄС. Однією з амбітних заяв стала можливість створення цілого міністерства під керівництвом ШІ [6, 13].

Прем'єр-міністр Еді Рама розглядає цю технологію, як хорошу можливість прозорої боротьби з корупцією. Міністр вважає штучний інтелект менш упередженим та вразливим за людей. Тож вбачає велику користь від нього у запобіганні «кумовству». Еді Рама також вважає, що одного дня розробники могли б створити справжню модель для обрання міністра. Додає, що тоді країна «стане першою, яка матиме цілий уряд з міністрами зі штучним інтелектом та прем'єр-міністром». Албанія доволі давно бореться з корупцією і вбачає великий шанс у розвитку цієї боротьби саме за допомогою технологій ШІ [6, 13].

Свій вартий уваги підхід має і Китай. У 2023 році Китай випустив регуляторний акт Interim Measures for the Management of Generative Artificial Intelligence Services, який наразі є чинним. Акт стосується генеративних систем ШІ, що спрямовані на генерування тексту, зображень, відео, наприклад, мовні моделі та інше. Однією з особливостей даного регуляторного документу є те, що він не розповсюджується на дослідження чи розробку генеративних моделей ШІ, які не призначені для публічного використання. Таким чином китайська адміністрація взяла під контроль загально-використовувані моделі, але не обмежила дослідження розвитку штучного інтелекту, як явища [14].

З одного боку Interim Measures for the Management of Generative AI Services заохочує використання генеративних моделей у різних сферах. Проте так само і

забороняє його використання з метою створення контенту, що підриває державну владу, неспокій у суспільстві, екстремізм, насильство, порнографія, неправдива інформація, таким чином окреслюючи певні «червоні лінії» [14].

Усі постачальники мають дотримуватись чинного законодавства. У разі виникнення проблем, таких як генерація шкідливого або небезпечного контенту, вони мають негайно зреагувати та виправити модель, видалити або замінити контент. Увага також приділяється сервісам, що можуть впливати на загальну думку. Для них проводяться безпекові оцінки та алгоритмічна звітність. Постачальники, що порушують усталені правила караються штрафами чи навіть позбавленням права вести діяльність [14].

White Paper: «A pro-innovation approach to AI regulation» – британський документ пов’язаний з регулюванням ШІ. Документ покликаний для підтримки розвитку штучного інтелекту, з урахуванням можливих ризиків та забезпеченням довіри громадян. Він ґрунтується на трьох основних засадах: пропорційність, інноваційність та контекст-чутливість. Завдяки пропорційності регулювання не є більш жорстким, ніж це потрібно. Замість того, щоб встановлювати однакові вимоги для всіх типів ШІ, як це робить ЄС, Велика Британія дозволяє різним сферам, наприклад, фінанси, медицина чи транспорт, самим визначати, коли і які стандарти потрібно застосувати [15].

За правилом інноваційності – усі інновації у даній галузі повинні сприяти загальному розвитку та економічному зростанню. Для цього британський уряд ухвалив створення спеціального середовища для тестування ШІ моделей – AI Regulatory Sandbo [15].

Контекст-чутливість має на увазі більшу довіру регуляторам, які відносяться до тієї чи іншої сфери. Тобто штучний інтелект у банківських алгоритмах, медичній структурі чи соціальних мережах – це різні середовища, з різними ризиками і можливими наслідками. Тож за документом вважається, що, наприклад, Ofcom – медіа, FCA – фінанси, знатимуть які краще регуляторні заходи вводити [15].

Канада проводить регулювання штучного інтелекту за допомогою Artificial Intelligence and Data Act (AIDA), який є частиною Закону Bill C-27. Акцент робиться на відповідальному та етичному використанні ШІ. Головна ідея AIDA – ризик-орієнтований підхід, що акцентує більшу увагу саме на системах, які можуть завдати шкоди здоров'ю, безпеці чи правам людини. Відтак, організації, що випускають моделі з подібними ризиками мусять проводити регулярну оцінку ризиків, запобігати упередженому ставленню та забезпечувати прозорість під людським контролем [16].

Також за документом створено спеціальний орган для контролю його роботи і виконання – AI and Data Commissioner. Порушення можуть тягти адміністративну або навіть кримінальну відповідальність. Наприклад, за використання незаконно отриманих даних для навчання моделей буде задіяна кримінальна відповідальність. Канада прагне поєднати довіру, безпеку та розвиток технологій, пропонуючи гнучке та зрозуміле регулювання [16].

У 2024 році Південна Корея також прийняла регуляторний акт з безпеки використання штучного інтелекту – Artificial Intelligence (AI) Basic Act. Він передбачає створення центральних механізмів управління: «контрольної вежі» – AI control tower та інституту безпеки ШІ – AI Safety Institute, які координуватимуть діяльність у сфері досліджень, стандартів, інфраструктури та підтримки стартапів [17].

Для бізнесу закон містить чіткі зобов'язання: компанії, які розробляють або використовують ШІ-системи, мають прямий вплив на людей, а також генеративні ШІ, повинні проводити оцінку ризиків, впроваджувати заходи забезпечення безпеки та прозорості, а також мати місцевого представника в Кореї. Установлено, що закон набуде чинності з січня 2026 року, а його підзаконні акти вже почали діяти з 2025 року [17].

2.2 Порівняльний аналіз підходів

Для чіткого розуміння підходів інших країн ми сформували таблицю із чіткими загальними характеристиками різних підходів (табл. 2.1).

Таблиця 2.1 – Загальна характеристика підходів

Країна / Регіон	Тип підходу	Рівень регулювання	Основна мета	Характер регулювання
ЄС (AI Act)	Комплексний, законодавчий	Високий	Захист прав людини, контроль ризиків	Жорстке правове регулювання, ризик-орієнтований підхід
США (NIST AI RMF)	Добровільний, управлінський	Середній	Управління ризиками, етичне використання	Добровільні рамки, інструменти саморегулювання
ISO/IEC 23894	Міжнародний стандарт	Низький	Системний підхід до менеджменту ризиків	Методичний, універсальний, придатний до адаптації
Велика Британія (White Paper)	Проінноваційний, гнучкий	Середній	Стимулювання розвитку технологій	Децентралізований, контекстно-чутливий підхід
ОАЕ	Галузевий, адаптивний	Середній	Стратегічна інтеграція ШІ у всі сектори	Публічно-приватні партнерства, експериментальні моделі
Канада (AIDA)	Законодавчий, етичний	Високий	Прозорість, підзвітність, безпечність	Регулювання через уповноважений орган
Південна Корея (AI Basic Act)	Централізований, безпековий	Високий	Контроль, стандартизація, інновації	Централізований орган контролю
Китай	Контрольований, ідеологічний	Високий	Соціальна стабільність, безпека	Обмеження генеративного контенту, жорсткий контроль
Албанія	Політичний, експериментальний	Низький	Демонстрація технологічної спроможності	Використання ШІ у державному управлінні

2.3 Теоретичний огляд можливості інтеграції міжнародних підходів в Україні

На основі проведеного теоретичного дослідження можна розглянути перспективність імплементації різних міжнародних досвідів в Україні.

Першими і важливим для інтеграції в Україні є саме європейські регуляторні акти та правові норми. Оскільки Україна активно націлена саме на інтеграцію з ЄС. AI Act [8] може бути основною точкою орієнтиру. Його модель побудована на оцінці ризиків і встановлює зрозумілу градацію: від неприйнятних до високих і низьких ризиків. Для України інтеграція цього підходу дасть змогу

створити систему, яка захищає громадян від небезпечних застосувань ШІ, але не блокує інновації.

США роблять акцент на добровільному, гнучкому регулюванні. NIST AI RMF [9] орієнтується на практичні інструменти управління ризиками (Govern, Map, Measure, Manage, Improve). Цей підхід вартий уваги для України, адже він дозволяє підприємствам відпрацьовувати внутрішні політики без надмірного тиску держави. Водночас інтеграція таких практик може підготувати бізнес до жорсткіших вимог ЄС.

ОАЕ [12] показують, що «гнучке регулювання» із залученням публічно-приватних партнерств може стати ефективним інструментом швидкої адаптації. Для України цей досвід цікавий тим, що дозволяє через галузеві стандарти (наприклад, у фінансах чи транспорті) створювати правила, навіть, без єдиного рамкового закону. Це корисно на перехідному етапі, коли AI Act ще не повністю імплементований.

Албанія [13] показує «політичний експериментальний» підхід – спроби інтегрувати ШІ, навіть, у державні інституції як символ модернізації. Для України цей досвід може слугувати радше застереженням: важливо уникати популізму й надмірного «олюднення» ШІ в управлінні, фокусуючись на практичних і етичних питаннях.

Схожий інтерес для України складає досвід Великої Британії [15]. Її підхід базується на принциповому регулюванні та делегуванні значної частини відповідальності галузевим регуляторам. Така модель сприяє швидкій адаптації норм під реалії конкретних секторів і стимулює інновації.

Для України британська модель може бути корисною на етапі формування спеціалізованих стандартів, коли необхідно балансувати між контролем і технологічним розвитком, орієнтуючись на потреби окремих сфер – оборони, медицини, транспорту тощо.

Канада у своєму законопроєкті AIDA [16] пропонує орієнтацію на прозорість, відповідальність розробників і пріоритет етичних засад. Її досвід

цікавий тим, що у своєму підході Канада поєднує принципи гнучкості з акцентом на соціальну відповідальність і права людини.

Для України інтеграція елементів цього підходу може сприяти формуванню гармонійної моделі, яка враховує потреби суспільства, бізнесу й держави. При цьому Канада демонструє приклад поступового та обережного впровадження регулювання, що може бути корисним у перехідний період в Україні.

Південна Корея [17] робить ставку на активну державну підтримку інновацій і водночас формує систему безпеки та етичного використання ШІ. Прийнятий AI Basic Act спрямований на стратегічний розвиток технологій, створення інфраструктури, освітніх програм і підтримку індустрії.

Для України цей досвід є цінним у контексті державної політики стимулювання інновацій, розвитку технічної освіти й забезпечення технологічного суверенітету. Така модель може допомогти Україні зміцнити власний науково-технологічний потенціал і розбудувати національну екосистему ШІ.

Китай [14], у свою чергу, концентрується на державному контролі, етичному нагляді та національній безпеці. Регуляція тут розглядається як інструмент гарантування суспільної стабільності та національних інтересів.

Хоча цей підхід менш придатний для прямої імплементації в Україні через відмінні політичні та правові принципи, його аналіз є корисним у контексті формування рішень щодо інформаційної безпеки і кіберзахисту.

Деякі механізми превентивного контролю ризиків та стандарти безпеки можуть бути адаптовані, зберігаючи при цьому демократичні цінності та принципи відкритості.

3 РЕКОМЕНДАЦІЇ ТА СТВОРЕННЯ МОДЕЛІ БЕЗПЕКИ ШІ ДЛЯ УКРАЇНИ

3.1 Огляд поточної ситуації в українському правовому полі

Україна сформувала поетапну дорожню карту з впровадження регулювання штучного інтелекту, орієнтовану на інтеграцію в європейський цифровий простір і поступове впровадження стандартів ЄС, зокрема положень AI Act. Головна мета цього документа – створення системи регулювання, яка поєднує захист прав людини, розвиток інноваційної економіки та підвищення міжнародної конкурентоспроможності українських компаній [18-20].

Запропонований підхід ґрунтується на принципі bottom-up, тобто поступового переходу від «м’яких» регуляторних механізмів до формального законодавчого врегулювання. На першому етапі (розрахованому на 2-3 роки) планується створити сприятливе регуляторне середовище – тестові платформи, добровільні кодекси поведінки, механізми оцінки ризиків і публікацію «Білої книги», що міститиме рекомендації для держави та бізнесу. Це дасть змогу учасникам ринку поступово адаптуватися до майбутніх вимог без надмірного адміністративного тиску [18-20].

Другий етап передбачає поетапне впровадження норм європейського AI Act, особливо в контексті інтеграції України до Європейського Союзу. У межах цього етапу планується адаптація найвищих стандартів у сфері захисту прав людини, прозорості та етики використання ШІ, із урахуванням особливостей національного ринку. Такий підхід дозволить Україні не лише відповідати міжнародним вимогам, а й зберегти гнучкість та конкурентоспроможність на світовій арені [18-20].

У «Білій книзі» визначено три стратегічні цілі: захист прав людини, підтримка конкурентоспроможності бізнесу та євроінтеграція. Україна прагне гармонізувати майбутнє законодавство з правовими нормами ЄС, щоб забезпечити українським розробникам штучного інтелекту вільний доступ до європейського ринку. Водночас у сфері оборони регулювання ШІ наразі не

передбачається, з огляду на військовий стан і потребу у швидкому технологічному розвитку для захисту держави [18, 19].

Особливу увагу в документі приділено балансу між правами людини та інноваційністю. Міністерство цифрової трансформації наголошує, що надмірне регулювання може стримувати розвиток індустрії, а його повна відсутність – створити ризики для громадянських свобод. Тому Україна орієнтується на сервісну модель держави, яка сприяє розвитку ШІ через підтримку та надання зручних інструментів. Серед них – публічний веб-портал відповідального ШІ, платформа юридичної допомоги та інструмент добровільного маркування систем [18, 19].

Ключовим елементом регулювання стане методологія оцінки впливу ШІ на права людини. Вона допоможе визначати рівень ризику конкретного продукту та стане обов’язковою умовою для участі в регуляторній пісочниці або отримання державних консультацій. Така пісочниця дозволить стартапам і компаніям тестувати свої рішення під наглядом держави, готуючись до майбутніх стандартів. Додатково планується впровадження системи добровільного маркування ШІ, подібної до «етикеток якості» – щоб користувачі могли розуміти принципи роботи системи, оцінювати її надійність, безпечність та етичність [18, 19].

3.2 Модель безпеки ШІ в Україні

Одним із ключових елементів регулювання штучного інтелекту є побудова системи безпеки, яка гарантує відповідальне, етичне та надійне використання ШІ. Для України така модель повинна враховувати як світові стандарти, так і національні виклики, пов’язані з війною, цифровою трансформацією та інтеграцією в європейський правовий простір. У цьому розділі подано узагальнений підхід до формування моделі безпеки використання ШІ в українських умовах, модель наведено на рис. 3.1 [18].

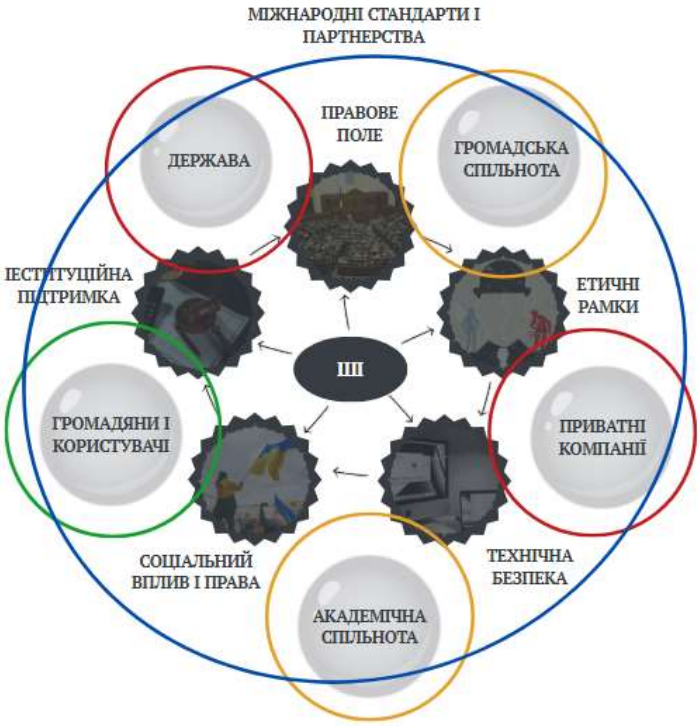

Рисунок 3.1 – Модель регулювання безпеки штучного інтелекту в Україні

Запропонована модель безпеки використання штучного інтелекту базується на системному підході, який передбачає взаємодію різних зацікавлених сторін і ключових сфер впливу. У центрі моделі знаходиться ШІ як технологічне ядро, довкола якого формуються міжсекторальні механізми регулювання, етики та технічної підтримки [18].

У моделі закладено взаємодію між державою, академічною спільнотою, громадянським суспільством, приватними компаніями, міжнародними партнерами та самими користувачами. Такий формат дозволяє не лише регулювати ШІ, а й формувати довіру до нього як до інструменту для розвитку, а не загрози. Україна – країна, що веде активну цифровізацію і не боїться впроваджувати нові технології. У цьому плані український громадський сектор та законодавча база є більш гнучкою і готовою до змін, аніж у сусідніх країнах Європейського Союзу [18].

Відповідно до рівня ризику який може виникнути у тій чи іншій складовій її позначено відповідним кольором (де, зелений – допустимий, жовтий – високий, червоний – підвищений). За концепцією схожою з AI Act цей поділ є необхідним

для оцінки кожної моделі ШІ, яка використовується у тій чи іншій галузі. Україна могла б створити «відкритий етичний реєстр ШІ», куди кожна компанія добровільно додала б свою модель. І, де кожна модель має власний «паспорт прозорості» – хто її створив, як навчав, які ризики враховані. Додатково кожна система проходила б оцінку на рівень ризику, відповідно до цього її власники мали б забезпечувати певний рівень безпеки, який відповідав би українському стандарту. Це не цензура, це – цифрова культура, адже користувачі мають знати, що саме використовують. Відповідно в уряді має бути створено спеціальний департамент з питань штучного інтелекту, який видавав би ліцензії безпеки та розглядав би і приймав усі потрібні нормативно-правові рішення. Також одним із важливих рішень має бути обов'язкове маркування продуктів та контенту створеного штучним інтелектом, задля запобігання шахрайства та недоброчесності [18].

Окрему роль у цій моделі може відігравати громадська спільнота, залучена до моніторингу та аудиту ШІ-систем. Запровадження публічних механізмів контролю, таких як цифрові платформи для фіксації порушень або несправедливих рішень, дозволить забезпечити не лише технічну, а й соціальну безпеку. Водночас державні сервіси – зокрема платформа «Дія» – можуть стати прикладом прозорого, етичного використання ШІ в публічному управлінні, де кожен громадянин має доступ до зрозумілих пояснень рішень, прийнятих алгоритмом [18].

Ключову роль також відіграє академічна спільнота, яка здатна не лише досліджувати ризики, а й формувати освітню культуру довкола ШІ. Впровадження національного освітнього треку з етики та технологій, інтеграція відповідних курсів у навчальні програми та підтримка молодих дослідників допоможе створити критичну масу спеціалістів, здатних формувати відповідальне майбутнє ШІ в Україні [18].

В українському контексті приватні компанії не обмежуються лише впровадженням технологій – вони стають співавторами національної цифрової безпеки. Особливо у сфері штучного інтелекту бізнес має не лише комерційні, а й

моральні зобов’язання: перед клієнтами, державою, суспільством і, навіть, перед власними працівниками. Українські компанії вже довели, що можуть бути лідерами в етичному програмуванні, відкритих кодах, благодійних проєктах та безпечних цифрових рішеннях [18].

Майбутнє моделі безпеки ШІ в Україні передбачає особливий соціальний договір між державою та бізнесом: не через штрафи чи контроль, а через довіру, сертифікацію, кооперацію та інноваційне партнерство. Компанії, які добровільно дотримуються стандартів прозорості, маркують свої моделі, відкривають алгоритми на ревізію, можуть отримати «цифрову довіру» – умовне етичне маркування, схоже на ISO, але у сфері штучного інтелекту. Водночас важливо підтримати компанії не тільки вимогами, а й ресурсами та середовищем. Наприклад, держава може створити інкубатори відповідального ШІ, де стартапи отримуватимуть доступ до відкритих даних, консультацій з етики, шаблонів політик – без тиску, без формальностей, з орієнтацією на співпрацю [18].

Ще однією важливою частиною запропонованої моделі безпеки використання ШІ є відкритість до міжнародної співпраці. Це включатиме обмін даними, впровадження міжнародних стандартизацій та регуляторних актів у інтегрованому для України форматі [18].

У контексті безпеки штучного інтелекту особливу увагу слід приділяти технічним практикам розробки та впровадження ШІ-рішень у приватному секторі, зокрема в компаніях, які створюють системи для використання в соціально чутливих сферах: оборона, фінанси, охорона здоров’я, освіта. Український бізнес сьогодні має справу не лише з комерційними викликами, а й із загрозами кібератак, викрадення моделей, отруєння даних та спробами маніпулювання алгоритмами. Ключовими напрямами технічної безпеки в межах української моделі ШІ можуть стати: безпечне навчання моделей, моніторинг поведінки моделей у реальному часі, розмежування доступу до компонентів ШІ систем, впровадження explainable AI та контейнери для тестування [18].

Окремої уваги заслуговує і питання інфраструктурної стійкості: більшість українських ІТ-компаній розміщують свої сервіси в хмарі, часто – за кордоном. В

умовах воєнних ризиків важливо створити резервні дата-центри, енергонезалежні вузли або використовувати «розподілену відповідальність» за безпеку з партнерами по хмарним сервісам, із обов'язковим аудитом [18].

Таким чином запропонована модель безпеки використання ШІ надаватиме простір для розвитку та новаторства, але при цьому увесь процес від створення моделей штучного інтелекту до їх подальшого безпечного використання буде контрольованим та безпечним. Відкритість до міжнародної співпраці тільки підтримує намір України підтримувати європейські стандарти безпеки та при цьому дасть можливість ділитись власним досвідом з міжнародними партнерами [18].

3.3 Рекомендації та застереження для України

Враховуючи вищезазначену модель безпеки використання штучного інтелекту в Україні, а також наявний міжнародний досвід інтеграції цих технологій, можна виділити наступні рекомендації для України.

Перше і доречне – визначення сфер у яких використання ШІ буде заборонено або обмежено. Важливо чітко окреслити межі використання штучного інтелекту. У наших рекомендаціях ми виділили такі сфери для обмеження: державний апарат, безпека та оборона, медицина та інші критичні сфери. Проте винятки також мають бути прописані окремо. Заборонити використання ШІ варто у застосунках та іграшках для дітей, адже це є однією із вразливих груп людей. Запобігання ризику маніпулювання та спонуканню до небезпечних дій є важливим аспектом регулювання. Також використання ШІ є неприйнятним у роботі з вразливими даними і персональною інформацією.

Моделі штучного інтелекту також мають обмежуватись та детально перевірятись у випадку використання, як медичні радники чи психологічна підтримка. Тобто необхідно слідкувати та перевіряти чи модель не підтримує галюцинації, думки про самогубство та інші небезпечні дії.

Ще однією сферою, у якій не варто опиратись на технології штучного інтелекту, є політика. Враховуючи плани ОАЄ та Албанії можна помітити тенденцію вважати штучний інтелект неупередженим, а отже прийнятним для

використання у прозорому виборі посадовців чи запобіганню корупції. Моделі штучного інтелекту надають відповіді на основі наданих їм даних. При цьому нема гарантій, що стороння особа не втрутиться до цього процесу, надаючи заздалегідь неправдиві факти, чим може спотворити ту саму неупередженість та прозорість, до яких прагнуть у політичній сфері.

Створення державного департаменту з питань ШІ не суперечить чинному законодавству України. Його створення навпаки може облегшити регулювання, адже певний орган буде відповідати за розгляд питань, пов'язаних зі штучним інтелектом та взаємодіяти з іншими органами та міністерствами, у випадку суміжних чи суперечливих питань.

Добровільний етичний реєстр штучного інтелекту може стати, як державною, так і громадською ініціативою.

Ще однією важливою рекомендацією є введення обов'язкового маркування продуктів медіа та контенту з використанням ШІ. Оскільки штучний інтелект все частіше використовується у створенні контенту, реклами та інших візуальних продуктів, важливим кроком є захист громадян від введення в оману. Такі зміни можна реалізувати завдяки змінам Закону «Про рекламу» [21] чи «Про захист прав споживачів» [22]. Погодивши введення обов'язкового маркування матеріалів створених штучним інтелектом саме з метою публічного розповсюдження.

Абсолютно легальним механізмом, який є не менш важливим для законного регулювання потоку моделей штучного інтелекту є введення паспорту прозорості ШІ та обов'язкової оцінки. Підставою є Закон «Про стандартизацію» [23], що дозволяє добровільну стандартизацію технологічних процесів.

Та є й рекомендації, що потребують додаткових дій з боку держави. Так, наприклад, ліцензування систем ШІ, поки не передбачувано законодавством. Це стане можливим тільки після ухвалення спеціального закону щодо штучного інтелекту, який визначить рамки та параметри ліцензування.

Відкриття алгоритмів для ревізії має бути обговорене та проводитись за протоколами, що не змушуватимуть розробників відкривати частини алгоритмів, які поставлять під ризик авторське право. Опорними законами у цьому питанні є

Закон «Про охорону прав на винаходи та корисні моделі» [24] і Цивільний кодекс [25], які не дозволяють примусово розкривати вихідний код без спеціального правового механізму.

Суперечливою у реалізації, але важливою є рекомендація щодо контролю моделей від створення до їх використання. Організація даного процесу має не суперечити ст. 42 Конституції України [26] – свобода підприємницької діяльності. Тож контроль має бути не адміністративним, а більш етичним, включати аудит та сертифікації.

3.4 Оцінка адекватності застосування міжнародних стандартів та регуляторних актів для України

З метою проведення оцінювання можливостей інтеграції досвіду інших країн у галузі штучного інтелекту для України, ми провели оцінку адекватності кожного з підходів у рамках українського простору.

3.4.1 Вибір методу оцінювання та програмно-апаратного забезпечення

Ми обрали удосконалену методику оцінювання для порівняння різних підходів за певними обраними критеріями. Адже дана методика є універсальною і може застосовуватись у різних випадках і для різних методів чи систем [27, 28].

Необхідними даними для проведення оцінки є обрані критерії, на основі яких буде проведено оцінку експертів, а після отримано значення для порівняння [27, 28].

Методика ґрунтується на експертному оцінювання за обраними критеріями. Для вибору експертів було враховано їх компетентність у даній галузі. В якості вихідних даних отримуються певні дані, що знаходяться за визначеними математичними виразами. Спираючись на дані, отримані від експертів, спочатку здійснюється математична обробка відповідно до правил, визначених у кожному методі оцінювання, а далі – на основі отриманих результатів проводиться програмне моделювання [27, 28].

В якості середовища для проведення розрахунків та обробки даних було обрано MathCad. Адже саме це середовище містить необхідним пакет інструментів для роботи.

3.4.2 Опис послідовності дій для оцінювання

Для проведення оцінювання було виконано наступні дії [27, 28]:

1) Вибір експертів за рівнем їх компетентності;
2) Вибір критеріїв (умовних та безумовних) для оцінювання;
3) Проведення оцінки за безумовними критеріями;
4) На підставі безумовного інтегрального критерію ухвалюється рішення – залежно від отриманої оцінки – щодо доцільності подальшого порівняння та аналізу;
5) За умови отримання позитивної оцінки – проведення оцінювання за умовними критеріями;
6) Опис результатів та висновки.

3.4.3 Опис та результати дослідження за сукупністю безумовних критеріїв

Для оцінки було обрано наступні безумовні критерії:

- K_{s1} – валідність;
- K_{s2} – гнучкість;
- K_{s3} – етичність;
- K_{s4} – складність;
- K_{s5} – прогресивність.

Під валідністю мається на увазі реалізованість методу (регуляторного акту чи стандарту): чи вже впроваджено на практиці, чи тільки заплановано на майбутнє.

Гнучкість – критерій, що описує інтегрованість підходу (регуляторного акту чи стандарту) у різні сфери.

Етичність характеризує те, чи метод справляється із захистом прав людини і чи не порушує підхід усталені норми етики та моралі.

Складність характеризує наявність ускладнень при інтеграції у іншій країні, в нашому контексті – в Україні.

Прогресивність – критерій для визначення того чи сприяє метод нормальному розвитку технології.

Оскільки перелічені часткові критерії належать до безумовних, відбір здійснюється за допомогою бінарної логічної змінної, а саме «так» чи «ні», або 1/0. Таким чином, безумовний критерій може бути представлений у формі (3.1) [27, 28]:

$$(K_{s1}, K_{s2}, K_{s3}, K_{s4}, K_{s5}) \in (1,0). \quad (3.1)$$

Враховуючи перераховані безумовні критерії та (3.1), функція відповідності адекватності методу (регуляторного акту чи стандарту) може бути записана наступним чином (3.2) [27, 28]:

$$f_{\text{фв}}() = (K_{s1} \wedge K_{s2} \wedge K_{s3} \wedge K_{s4} \wedge K_{s5}) = K_s. \quad (3.2)$$

Отже, якщо $f_{\text{фв}} = 1$ – значить метод відповідає вимогам. Ті методи, що відповідають вимогам будуть розглянуті далі.

Результати порівняння відносно безумовних критеріїв наведено у табл. 3.1.

Таблиця 3.1 – Результати порівняння відносно безумовних критеріїв

Критерій / Метод	K_{s1}	K_{s2}	K_{s3}	K_{s4}	K_{s5}	K_s
ЄС (AI Act)	1	1	1	1	1	1
США (NIST AI RMF)	1	1	1	1	1	1
ISO/IEC 23894	1	1	1	1	1	1
Велика Британія (White Paper)	1	1	1	1	1	1
ОАЕ	0	1	0	0	1	0
Канада (AIDA)	1	1	1	1	1	1
Південна Корея (AI Basic Act)	1	1	1	1	1	1
Китай	1	1	1	1	1	1
Албанія	0	1	0	0	1	0

Отже, негативну оцінку отримали методи запропоновані у Албанії та ОАЕ [12, 13]. Усі інші будуть розглянуті у подальших дослідженнях на основі умовних критеріїв та інтегрального умовного критерію.

3.4.4 Опис та результати дослідження за сукупністю умовних критеріїв

В якості умовних критеріїв було обрано часткові критерії наведені у табл. 3.2. Для порівняння підходів було побудовано дерево, зображене на рис. 3.2. Наступним кроком було побудовано табл. 3.3, в якій наведено оцінку вкладу кожного критерію.

Таблиця 3.2 – Умовні часткові критерії оцінки ЕП

№	Критерії	Позначення
1	Рівень прозорості алгоритмів і систем прийняття рішень, тобто пояснюваність процедур, вимог та інших дій зазначених у документі.	K_{y1}
2	Перспективність інтеграції регуляторного акту, методу, стандарту в Україні.	K_{y2}
3	Рівень гнучкості документу у застосуванні в різних сферах.	K_{y3}
4	Можливість запобігання виникненню етичних ризиків, пов'язаних з порушенням прав та свобод людини.	K_{y4}
5	Рівень інституційної підтримки та державної залученості.	K_{y5}
6	Вплив на перспективність розвитку технології окремо та у інших технологічних галузях.	K_{y6}

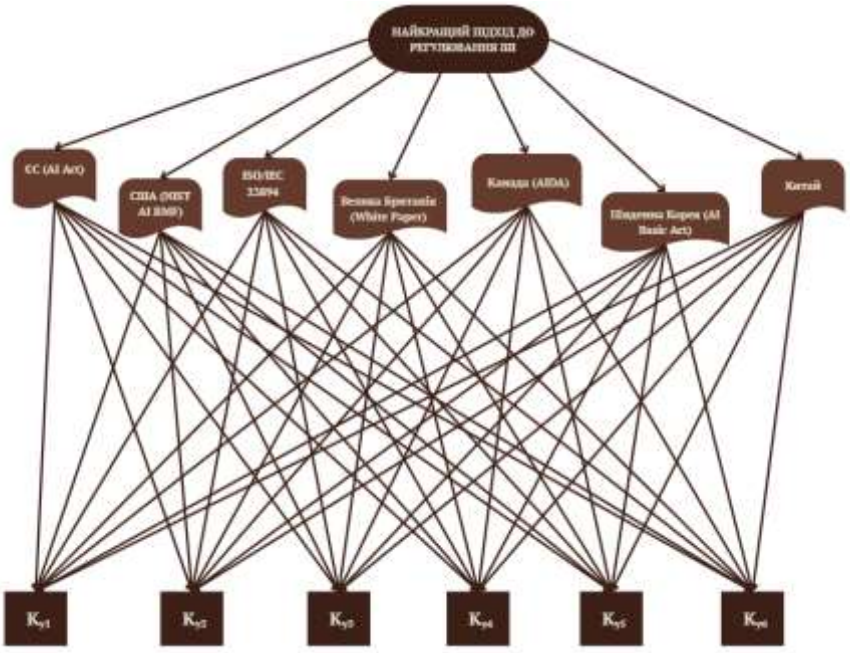

Рисунок 3.2 – Дерево порівняння

Для оцінки кожного критерію було побудовано матрицю попарних порівнянь.

Таблиця 3.3 – Попарне порівняння критеріїв для визначення важливості

Критерії	K_{y1}	K_{y2}	K_{y3}	K_{y4}	K_{y5}	K_{y6}	q_j	r_j
K_{y1}	1	1	2	0,5	4	2	1,41421356	0,18940253
K_{y2}	1	1	1	0,33	2	3	1,12246205	0,15032889
K_{y3}	0,5	1	1	0,2	3	1	0,81818882	0,10957824
K_{y4}	2	3	6	1	6	3	2,94168275	0,39397315

Подовження таблиці 3.3.

Критерії	K_{y1}	K_{y2}	K_{y3}	K_{y4}	K_{y5}	K_{y6}	q_j	r_j
K_{y5}	0,25	0,5	0,33	0,1667	1	0,1428	0,31580809	0,04229549
K_{y6}	0,5	0,33	1	0,33	7	1	0,85435347	0,1144217

Відношення узгодженості має значення 7,466708745.

Таким чином ми отримали розуміння того, який з критеріїв буде відігравати найбільшу роль при оцінюванні.

Під час парного порівняння експерт оцінює об’єкти попарно, визначаючи, який із двох є важливішим. Усі можливі комбінації об’єктів подаються у вигляді матриці попарних порівнянь. Під час проведення парних порівнянь необхідно відповісти на запитання: який із двох елементів є важливішим, має більший вплив, вищу ймовірність або перевагу. Зазвичай при порівнянні критеріїв визначають, який із них є значущішим відносно один одного [27, 29].

Отримані значення показують, що серед усіх критеріїв найбільшу вагу має етичний компонент K_{y4}, який відіграє ключову роль у побудові безпечних і прозорих систем ШІ. Це свідчить про те, що під час подальшого аналізу саме цей критерій найбільше впливатиме на результат оцінювання.

Таким чином було проведено порівняння окремо за критеріями табл. 3.4-3.9.

Таблиця 3.4 – Матриця попарних порівнянь за критерієм K_{y1}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	1	1	2	2	2	3	1,574610106	0,20586303
США (NIST AI RMF)	1	1	1	2	2	2	3	1,574610106	0,20586303
ISO/IEC 23894	1	1	1	2	2	2	3	1,574610106	0,20586303
Велика Британія (White Paper)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Канада (AIDA)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Південна Корея (AI Basic Act)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Китай	0,33333333	0,333333333	0,333333333	0,5	0,5	0,5	1	0,463987849	0,06066133

Відношення узгодженості для першого критерію дорівнює 7,648824235.

Результати попарного порівняння за критерієм K_{y1} демонструють перевагу підходів, що ґрунтуються на комплексному регулюванні.

Таблиця 3.5 – Матриця попарних порівнянь за критерієм K_{y2}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	2	1	1	2	2	7	1,777219815	0,21271988
США (NIST AI RMF)	0,5	1	0,5	0,5	1	1	5	0,935061127	0,1119198
ISO/IEC 23894	1	2	1	1	2	2	7	1,777219815	0,21271988
Велика Британія (White Paper)	1	2	1	1	2	2	7	1,777219815	0,21271988
Канада (AIDA)	0,5	1	0,5	0,5	1	1	5	0,935061127	0,1119198
Південна Корея (AI Basic Act)	0,5	1	0,5	0,5	1	1	5	0,935061127	0,1119198
Китай	0,14285714	0,2	0,142857143	0,142857143	0,2	0,2	1	0,21789966	0,02608095

Відношення узгодженості для другого критерію дорівнює 8,354742485.

Як видно з таблиці 3.5, за критерієм K_{y2} найкращі результати показали ЄС, ISO/IEC та Велика Британія, що свідчить про можливість адаптації їхніх моделей до українського правового поля. Порівняно з ними, китайський підхід має суттєво нижчий потенціал інтеграції через відмінності в законодавчих принципах.

Таблиця 3.6 – Матриця попарних порівнянь за критерієм K_{y3}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	0,5	0,33333333	0,5	1	1	2	0,7741686	0,0967747
США (NIST AI RMF)	2	1	0,5	1	2	2	3	1,4261616	0,17827688
ISO/IEC 23894	3	2	1	2	3	3	4	2,3795656	0,29745684
Велика Британія (White Paper)	2	1	0,5	1	2	2	3	1,4261616	0,17827688
Канада (AIDA)	1	0,5	0,33333333	0,5	1	1	2	0,7741686	0,0967747
Південна Корея (AI Basic Act)	1	0,5	0,33333333	0,5	1	1	2	0,7741686	0,0967747
Китай	0,5	0,33333333	0,25	0,33333333	0,5	0,5	1	0,4453057	0,05566529

Відношення узгодженості для третього критерію дорівнює 7,99970022.

Під час оцінки за критерієм K_{y3} найвищі результати отримав міжнародний стандарт ISO/IEC 23894, що демонструє його універсальність та придатність для різних галузей. Підходи США і Великої Британії мають подібні показники, підтверджуючи збалансований характер їхніх моделей регулювання.

Таблиця 3.7 – Матриця попарних порівнянь за критерієм K_{y4}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	2	2	3	3	4	5	2,5597076	0,30683667
США (NIST AI RMF)	0,5	1	1	2	2	3	4	1,5746101	0,18875129
ISO/IEC 23894	0,5	1	1	0,5	1	2	3	1,059634	0,1270202
Велика Британія (White Paper)	0,33333333	0,5	2	1	2	3	4	1,3459002	0,16133543
Канада (AIDA)	0,33333333	0,5	1	0,5	1	2	3	0,9057237	0,1085707
Південна Корея (AI Basic Act)	0,25	0,33333333	0,5	0,33333333	0,5	1	2	0,5428337	0,06507043
Китай	0,2	0,25	0,33333333	0,25	0,3333333	0,5	1	0,3538387	0,04241527

Відношення узгодженості для четвертого критерію дорівнює 8,342247982.

Як свідчать результати для K_{y4}, найбільшу вагу має європейський підхід, що узгоджується з високими вимогами до етичності й захисту прав людини в AI Act. Порівняно з ним, китайська модель суттєво поступається за цим показником, оскільки зосереджена передусім на контролі й безпеці, а не на гуманістичних принципах.

Таблиця 3.8 – Матриця попарних порівнянь за критерієм K_{y5}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	4	6	3	3	2	0,5	2,1552289	0,24590539
США (NIST AI RMF)	0,25	1	2	0,5	0,5	0,3333333	0,25	0,5209768	0,05944195

Подовження таблиці 3.8.

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ISO/IEC 23894	0,16666667	0,5	1	0,3333333	0,3333333	0,25	0,1666667	0,3253403	0,03712039
Велика Британія (White Paper)	0,33333333	2	3	1	1	0,5	0,3333333	0,8547514	0,09752467
Канада (AIDA)	0,33333333	2	3	1	1	0,5	0,3333333	0,8547514	0,09752467
Південна Корея (AI Basic Act)	0,5	3	4	2	2	1	0,5	1,4261616	0,16272092
Китай	2	4	6	3	3	2	1	2,6272534	0,29976202

Відношення узгодженості для п’ятого критерію дорівнює 8,764463803.

Оцінювання за критерієм K_{y5} показало, що ЄС і Китай демонструють найвищі результати, однак за різною природою: перший – завдяки міждержавним регуляторним інституціям, другий – через централізований урядовий контроль.

Таблиця 3.9 – Матриця попарних порівнянь за критерієм K_{y6}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_j	r_j
ЄС (AI Act)	1	0,5	4	0,5	1	0,33333333	4	1,0419536	0,12185127
США (NIST AI RMF)	2	1	3	1	2	0,5	3	1,5112094	0,17672839
ISO/IEC 23894	0,25	0,33333333	1	0,33333333	0,25	0,2	0,2	0,3104251	0,03630266
Велика Британія (White Paper)	2	1	3	1	2	0,5	3	1,5112094	0,17672839
Канада (AIDA)	1	0,5	4	0,5	1	0,33333333	4	1,0419536	0,12185127
Південна Корея (AI Basic Act)	3	2	5	2	3	1	5	2,6426196	0,30904115
Китай	0,25	0,33333333	5	0,33333333	0,25	0,2	1	0,4916573	0,05749687

Відношення узгодженості для шостого критерію дорівнює 8,551027969.

Результати оцінки за критерієм K_{y6} вказують, що найбільш інноваційно орієнтованими є підходи США та Південної Кореї, які поєднують регуляцію з підтримкою досліджень і розробок. Це контрастує з китайським підходом, який

демонструє нижчі показники через переважання контрольних механізмів над стимулюючими.

Розрахунки вектору пріоритетів було проведено на основі добутку вектору пріоритетів першого рівня, отриманого раніше у табл. 3.3, та матриці отриманих значень.

$$t := (0.1840253\ \ 0.15032889\ \ 0.10957824\ \ 0.39397315\ \ 0.04229549\ \ 0.1144217)$$

$$L := \begin{pmatrix} 0.20586303 & 0.20586303 & 0.20586303 & 0.10724986 & 0.10724986 & 0.10724986 & 0.06066133 \\ 0.21271988 & 0.1119198 & 0.21271988 & 0.21271988 & 0.1119198 & 0.1119198 & 0.02608095 \\ 0.0967747 & 0.178227688 & 0.29745684 & 0.178227688 & 0.0967747 & 0.0967747 & 0.05566529 \\ 0.3063667 & 0.18875129 & 0.1270202 & 0.16133543 & 0.1085707 & 0.06507043 & 0.04241527 \\ 0.24590539 & 0.059441995 & 0.03712039 & 0.09752467 & 0.09752467 & 0.16272092 & 0.29976202 \\ 0.1218512 & 0.17672839 & 0.03630266 & 0.17672839 & 0.12185127 & 0.30904115 & 0.05749687 \end{pmatrix}$$

$$tl := t \cdot L = (0.226\ \ 0.171\ \ 0.158\ \ 0.159\ \ 0.108\ \ 0.115\ \ 0.057)$$

Рисунок 3.3 – Розрахунок результуючого вектору пріоритетів

У результаті отримано чисельні дані за якими найкращі результат показав європейський підхід – ЄС (AI Act) – 0,226, а найгірші значення у контексті обраних критеріїв отримав китайський підхід до регулювання ШІ – 0,057. Значення інших розглянутих підходів налічують: США (NIST AI RMF) – 0,171; ISO/IEC 23894 – 0,158; Велика Британія (White Paper) – 0,159; Канада (AIDA) – 0,108; Південна Корея (AI Basic Act) – 0,115.

3.4.5 Оцінювання адекватності застосування міжнародних методів методом визначення вагових коефіцієнтів за допомогою шкали Фішберна

Даний метод полягає у тому, що за основу беруться оцінки певної кількості експертів, у нашому випадку їх було п’ять. Експерти оцінювали важливість того чи іншого критерію ранжуючи їх [27, 28, 30].

Для початку було проведено ранжування експертами визначених критеріїв (табл. 3.10).

Таблиця 3.10 – Ранжування критеріїв експертами

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	2	1	4	3	5	6
2	5	1	3	4	2	6
3	3	2	4	2	6	5

Подовження таблиці 3.10.

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
4	5	1	4	2	6	3
5	3	2	5	1	6	4

Після було проведено розрахунки для визначення значень вагових коефіцієнтів за формулою [27, 28, 30]:

$$a_i = \frac{2\cdot(m-i+1)}{m\cdot(m+1)},$$

де a_i – ваги i-го показника, i – номер показника, m – кількість показників.

Результати визначення значень вагових коефіцієнтів для критеріїв наведено у табл. 3.11.

Таблиця 3.11 – Значення вагових коефіцієнтів та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,23809524	0,28571429	0,14285714	0,19047619	0,095238	0,047619
2	0,0952381	0,28571429	0,19047619	0,14285714	0,238095	0,047619
3	0,19047619	0,23809524	0,14285714	0,28571429	0,047619	0,0952381
4	0,0952381	0,28571429	0,14285714	0,23809524	0,047619	0,1904761
5	0,19047619	0,23809524	0,0952381	0,28571429	0,047619	0,1428571
w_i	0,16190476	0,26666667	0,14285714	0,22857143	0,095238	0,1047619

а) Оцінювання підходу зазначеного у документі ЄС (AI Act)

Спочатку експерти проранжували критерії відповідно для документу ЄС (AI Act). Після було проведено розрахунок вагових коефіцієнтів та їх середнє значення – табл. 3.12.

Таблиця 3.12 – Вагові коефіцієнти для ЄС (AI Act) та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,238095238	0,28571429	0,19047619	0,14285714	0,047619	0,0952381
2	0,285714286	0,19047619	0,14285714	0,23809524	0,047619	0,0952381
3	0,19047619	0,23809524	0,0952381	0,28571429	0,047619	0,1428571
4	0,285714286	0,23809524	0,14285714	0,19047619	0,047619	0,0952381
5	0,238095238	0,28571429	0,0952381	0,19047619	0,142857	0,0476190
w_i	0,247619048	0,24761905	0,13333333	0,20952381	0,066666	0,0952381

b) Оцінювання підходу зазначеного у документі США (NIST AI RMF)

Експерти проранжували критерії для наступного учасника оцінювання – документу США (NIST AI RMF). Після було проведено розрахунок вагових коефіцієнтів та їх середнє значення – табл. 3.13.

Таблиця 3.13 – Вагові коефіцієнти для США (NIST AI RMF) та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,142857143	0,23809524	0,0952381	0,19047619	0,047619	0,2857142
2	0,095238095	0,28571429	0,14285714	0,04761905	0,238095	0,1904761
3	0,047619048	0,19047619	0,23809524	0,14285714	0,285714	0,0952381
4	0,19047619	0,0952381	0,04761905	0,23809524	0,142857	0,2857142
5	0,238095238	0,04761905	0,19047619	0,0952381	0,285714	0,1428571
w_i	0,142857143	0,17142857	0,14285714	0,14285714	0,2	0,2

c) Оцінювання підходу зазначеного у документі ISO/IEC 23894

Після експерти проранжували критерії для документу ISO/IEC 23894. Після чого за тим же принципом було розраховано значення вагових коефіцієнтів та їх середнє значення – табл. 3.14.

Таблиця 3.14 – Вагові коефіцієнти для ISO/IEC 23894 та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,285714286	0,19047619	0,0952381	0,23809524	0,047619	0,1428571
2	0,285714286	0,23809524	0,14285714	0,19047619	0,047619	0,0952381
3	0,238095238	0,19047619	0,04761905	0,28571429	0,095238	0,1428571
4	0,285714286	0,14285714	0,19047619	0,23809524	0,047619	0,0952381
5	0,238095238	0,19047619	0,0952381	0,28571429	0,047619	0,1428571
w_i	0,266666667	0,19047619	0,11428571	0,24761905	0,057142	0,1238095

d) Оцінювання підходу зазначеного у документі White Paper (Велика Британія)

Наступними експерти проранжували критерії відповідно для документу White Paper (Велика Британія). Після було знову проведено розрахунок вагових коефіцієнтів та їх середнє значення – табл. 3.15.

Таблиця 3.15 – Вагові коефіцієнти для White Paper (Велика Британія) та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,19047619	0,14285714	0,23809524	0,28571429	0,0952381	0,047619
2	0,238095238	0,19047619	0,28571429	0,14285714	0,0952381	0,047619
3	0,19047619	0,23809524	0,14285714	0,28571429	0,0952381	0,047619
4	0,142857143	0,19047619	0,23809524	0,28571429	0,0952381	0,047619
5	0,238095238	0,14285714	0,19047619	0,28571429	0,0952381	0,047619
w_i	0,2	0,18095238	0,21904762	0,25714286	0,0952381	0,047619

e) Оцінювання підходу зазначеного у документі AIDA (Канада)

Далі для документу AIDA (Канада) експерти проранжували критерії відповідно. Після було проведено розрахунок вагових коефіцієнтів та їх середнє значення – табл. 3.16.

Таблиця 3.16 – Вагові коефіцієнти для AIDA (Канада) та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,238095238	0,19047619	0,14285714	0,28571429	0,0952381	0,047619
2	0,19047619	0,23809524	0,14285714	0,28571429	0,0952381	0,047619
3	0,238095238	0,19047619	0,28571429	0,14285714	0,0952381	0,047619
4	0,19047619	0,23809524	0,28571429	0,14285714	0,0952381	0,047619
5	0,238095238	0,14285714	0,19047619	0,28571429	0,0952381	0,047619
w_i	0,219047619	0,2	0,20952381	0,22857143	0,0952381	0,047619

f) Оцінювання підходу зазначеного у документі AI Basic Act (Південна Корея)

Як у попередніх випадках експерти проранжували критерії відповідно для документу AI Basic Act (Південна Корея). Після було розраховано значення вагових коефіцієнтів та їх середнє значення – табл. 3.17.

Таблиця 3.17 – Вагові коефіцієнти для AI Basic Act (Південна Корея) та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,19047619	0,23809524	0,14285714	0,28571429	0,047619	0,0952381
2	0,238095238	0,19047619	0,14285714	0,28571429	0,047619	0,0952381
3	0,19047619	0,23809524	0,14285714	0,28571429	0,047619	0,0952381
4	0,238095238	0,14285714	0,19047619	0,28571429	0,047619	0,0952381
5	0,19047619	0,23809524	0,14285714	0,28571429	0,047619	0,0952381
w_i	0,20952381	0,20952381	0,15238095	0,28571429	0,047619	0,0952381

g) Оцінювання китайського підходу

Останніми експерти проранжували критерії відповідно для китайського підходу. Після було проведено розрахунок вагових коефіцієнтів та їх середнє значення – табл. 3.18.

Таблиця 3.18 – Вагові коефіцієнти для китайського підходу та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,285714286	0,14285714	0,0952381	0,19047619	0,238095	0,047619
2	0,19047619	0,14285714	0,04761905	0,28571429	0,238095	0,0952381
3	0,238095238	0,0952381	0,0952381	0,28571429	0,190476	0,047619
4	0,285714286	0,0952381	0,0952381	0,19047619	0,238095	0,047619
5	0,238095238	0,14285714	0,04761905	0,28571429	0,190476	0,0952381
w_i	0,247619048	0,12380952	0,07619048	0,24761905	0,219047	0,0666667

Надалі було проведено саме оцінювання та отримано результати – рис. 3.4. Найкращий результат отримав канадський підхід (AIDA) – 0,267. А найгірший результат – США (NIST AI RMF) – 0,162.

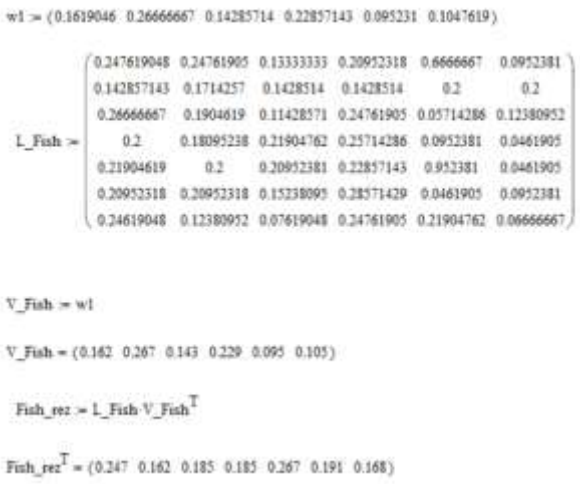

Рисунок 3.4 – Обчислення результуючого вектору пріоритетів

Інші підходи отримали наступні показники: ЄС (AI Act) – 0,247; китайський підхід – 0,168; ISO/IEC 23894 – 0,185; Велика Британія (White Paper) – 0,185; Південна Корея (AI Basic Act) – 0,191.

3.4.6 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі методу ранжування

Припустимо, що існує *m* часткових показників і *n* експертів, які оцінюють їхню важливість для певної системи [27, 28, 31]. У нашому випадку це знову ж п’ять експертів. Найвищий ранг (оцінку) *m* отримує показник, який вважається найважливішим, далі – показники з меншими рангами відповідно до зниження значущості, а ранг 1 присвоюється найменш важливому показнику. У цьому разі вагові коефіцієнти розраховуються за формулою [27, 28, 31]:

$$w_j = \frac{r_j}{\sum_{j=1}^{m} r_j}.$$

Як і у попередніх випадках, спершу було проведено розрахунки вагових коефіцієнтів для критеріїв – табл. 3.19.

Таблиця 3.19 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	6	5	2	4	3	1
2	5	4	1	6	3	2
3	6	4	3	5	2	1
4	5	3	2	6	4	1
5	6	4	2	5	1	3
$r_j = \sum_{i=1}^{n} r_{ij}$	28	20	10	26	13	8
w_j	0,26666667	0,19048	0,09524	0,24762	0,12381	0,07619

Також у наступних кроках було розраховано вагові коефіцієнти для кожного з підходів, які ми розглянули.

а) Оцінка для ЄС (AI Act)

Таблиця 3.20 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	5	4	3	6	2	1
2	4	5	2	6	3	1
3	5	4	3	6	1	2
4	4	3	2	6	5	1
5	5	4	3	6	2	1
$r_j = \sum_{i=1}^{n} r_{ij}$	23	20	13	30	13	6
w_j	0,21904762	0,19048	0,12381	0,28571	0,12381	0,05714

Отримані результати свідчать, що найбільшу вагу серед часткових показників для підходу ЄС (AI Act) має показник x_4 – 0,2857, що вказує на його ключову роль у загальній структурі оцінювання. Це може бути пов’язано з високим рівнем прозорості та етичності регуляторної системи Європейського Союзу, які є визначальними для впровадження принципів безпечного ШІ.

b) Оцінка для США (NIST AI RMF)

Таблиця 3.21 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	3	6	5	4	2	1
2	4	6	5	3	1	2
3	3	6	5	4	2	1
4	4	6	5	3	1	2
5	3	6	5	4	1	2
$r_j = \sum_{i=1}^{n} r_{ij}$	17	30	25	18	7	8
w_j	0,16190476	0,28571	0,2381	0,17143	0,06667	0,07619

Для підходу США (NIST AI RMF) спостерігається зміщення ваг у бік показника x_2 – 0,2857 – це відображає орієнтацію американської моделі на практичну гнучкість та стандартизацію процесів оцінювання ризиків. Порівняно з європейським AI Act, американський підхід демонструє більшу варіативність між окремими критеріями.

c) Оцінка для ISO/IEC 23894

Таблиця 3.22 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	6	4	5	3	2	1
2	6	4	5	3	1	2
3	6	5	4	3	2	1
4	6	5	4	2	3	1
5	6	4	5	3	2	1
$r_j = \sum_{i=1}^{n} r_{ij}$	30	22	23	14	10	6
w_j	0,28571429	0,20952	0,21905	0,13333	0,09524	0,05714

Як видно з табл. 3.22, для стандарту ISO/IEC 23894 найбільшу вагу отримали показники x_1 – 0,2857 і x_3 – 0,2191, що свідчить про його орієнтацію на

системність і структурованість підходів до управління ризиками штучного інтелекту. Отримані результати демонструють більш рівномірний розподіл вагових коефіцієнтів порівняно з американським підходом, що підкреслює міжнародну універсальність стандарту

d) Оцінка для White Paper (Велика Британія)

Результати, наведені у табл. 3.23, показують, що британський White Paper має дещо іншу логіку вагового розподілу. Найвищу вагу отримали показники x_2 – 0,2762 та x_3 – 0,2476, що вказує на пріоритетність критеріїв гнучкості та етичності у британському підході. Це відображає загальну філософію британської моделі регулювання ШІ – «pro-innovation approach», яка поєднує орієнтацію на інновації з мінімальним рівнем бюрократичних обмежень.

Таблиця 3.23 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	3	5	6	4	2	1
2	4	6	5	3	1	2
3	3	6	5	4	2	1
4	4	6	5	3	2	1
5	3	6	5	4	1	2
$r_j = \sum_{i=1}^{n} r_{ij}$	17	29	26	18	8	7
w_j	0,16190476	0,27619	0,24762	0,17143	0,07619	0,06667

e) Оцінка для канадського підходу (AIDA)

Таблиця 3.24 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	4	5	3	6	2	1
2	5	4	3	6	1	2
3	4	5	3	6	2	1
4	5	4	3	6	2	1
5	4	5	3	6	1	2
$r_j = \sum_{i=1}^{n} r_{ij}$	22	23	15	30	8	7
w_j	0,20952381	0,21905	0,14286	0,28571	0,07619	0,06667

Для канадського підходу AIDA найбільшу вагу має показник x_4 – 0,2857, який відображає акцент Канади на забезпеченні прозорості, етичності та

відповідальності при використанні ШІ. Це підтверджує прагнення уряду Канади сформувати збалансовану регуляторну систему, орієнтовану одночасно на захист прав людини та стимулювання інновацій. Показники мають доволі рівномірний розподіл, що вказує на комплексний характер підходу, де важливими є як інституційні, так і соціальні аспекти регулювання.

f) Оцінка для південнокорейського підходу (AI Basic Act)

У південнокорейському підході AI Basic Act домінують показники x_4 – 0,219 та x_5 – 0,2476, що відображає орієнтацію на державну підтримку, розвиток технологій та створення умов для інновацій. Південна Корея демонструє прагнення до формування стабільного законодавчого поля, однак із певним переважанням державного контролю, що дещо знижує гнучкість.

Таблиця 3.25 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	4	3	2	6	5	1
2	3	4	2	6	5	1
3	4	3	1	6	5	2
4	3	4	2	5	6	1
5	4	3	1	6	5	2
$r_j = \sum_{i=1}^{n} r_{ij}$	18	17	8	29	26	7
w_j	0,17142857	0,1619	0,07619	0,27619	0,24762	0,06667

g) Оцінка для китайського підходу

Для китайського підходу табл. 3.26 спостерігається значне зосередження ваг на показниках x_5 – 0,2762 та x_1 – 0,1809, що свідчить про централізований характер регулювання, орієнтований на контроль і державне управління розвитком технологій. Порівняно з попередніми підходами, китайська модель менш збалансована – частина критеріїв має суттєво нижчі значення, що вказує на обмежену гнучкість.

Таблиця 3.26 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	4	3	2	5	6	1
2	3	4	2	6	5	1
3	4	1	2	5	6	3
4	3	4	2	5	6	1
5	5	3	1	4	6	2
$r_j = \sum_{i=1}^{n} r_{ij}$	19	15	9	25	29	8
w_j	0,18095238	0,14286	0,08571	0,2381	0,27619	0,07619

$w2 := (0.26666667 \;\; 0.19048 \;\; 0.09524 \;\; 0.24762 \;\; 0.12381 \;\; 0.07619)$

$$L\_Koef := \begin{pmatrix} 0.21904762 & 0.19048 & 0.12381 & 0.28571 & 0.12381 & 0.05714 \\ 0.16190476 & 0.2851 & 0.2381 & 0.17143 & 0.06667 & 0.07619 \\ 0.28571429 & 0.20952 & 0.21905 & 0.13333 & 0.09524 & 0.05714 \\ 0.16190476 & 0.27619 & 0.24762 & 0.17143 & 0.07619 & 0.06667 \\ 0.20952381 & 0.21905 & 0.14286 & 0.2851 & 0.0619 & 0.06667 \\ 0.17142857 & 0.1619 & 0.0619 & 0.27619 & 0.24762 & 0.06667 \\ 0.18095238 & 0.14286 & 0.08571 & 0.2381 & 0.27619 & 0.07619 \end{pmatrix}$$

$V\_Koef := w2$

$V\_Koef = (0.267 \;\; 0.19 \;\; 0.095 \;\; 0.248 \;\; 0.124 \;\; 0.076)$

$Koef\_rez := L\_Koef \cdot V\_Koef^{T}$

$Koef\_rez^{T} = (0.197 \;\; 0.177 \;\; 0.186 \;\; 0.176 \;\; 0.195 \;\; 0.187 \;\; 0.183)$

Рисунок 3.5 – Обчислення результуючого вектору пріоритетів

Найкращий показник отримав європейський підхід – документ AI Act – 0,197. Найгірші результати показали Велика Британія (White Paper) – 0,176 та США (NIST AI RMF) – 0,177. Інші підходи отримали оцінки: ISO/IEC 23894 – 0,186; китайський підхід – 0,183; Південна Корея (AI Basic Act) – 0,187; Канада (AIDA) – 0,195.

3.4.7 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі методу приписування балів

У методі приписування балів маємо *m* показників, важливість яких оцінює *n* експертів. Кожен експерт визначає, наскільки важливим є кожен показник для досліджуваної системи, і виставляє йому оцінку за шкалою від 0 до 10 [27, 28].

Допускається використання дробових значень, якщо експерт вважає, що показник заслуговує проміжної оцінки. Також кілька показників можуть отримати однаковий бал, якщо їх важливість вважається рівною [27, 28].

Наступним кроком розраховується вага кожного коефіцієнта, використовуючи формули [27, 28]:

$$r_{ij} = \frac{h_{ij}}{\sum_{j=1}^{m} h_{ij}}, \quad \sum_{j=1}^{m} r_{ij} = 1,$$

де r_{ij} – ваги *j*-го показника, визначені *i*-м експертом; h_{ij}– бал *i*-го експерта, виставлений *j*-му показнику, *n* – кількість експертів, *m* – кількість показників.

В результаті вагові коефіцієнти визначались за формулами [27, 28]:

$$w_j = \frac{\sum_{i=1}^{n} r_{ij}}{\sum_{j=1}^{m}\sum_{i=1}^{n} r_{ij}}; \quad \sum_{j=1}^{n} w_j = 1.$$

Розрахунок вагових коефіцієнтів для критеріїв наведено у табл. 3.27.

Таблиця 3.27 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	8	5	10	7	6	45	0,2	0,1777	0,1111	0,2222	0,1555	0,1333
2	7	10	5	9	8	6	45	0,1555	0,2222	0,1111	0,2	0,1777	0,1333
3	6	7	9	8	5	10	45	0,1333	0,1555	0,2	0,1777	0,1111	0,2222
4	8	6	4	10	7	5	40	0,2	0,15	0,1	0,25	0,175	0,125
5	6	9	5	10	8	7	45	0,1333	0,2	0,1111	0,2222	0,17778	0,15556
							$\sum_{i=1}^{n} r_j$	0,8222	0,90555	0,6333	1,0722	0,7972	0,76944
							w_j	0,1644	0,1811	0,1266	0,2144	0,1594	0,15389

Наступним кроком було проведено розрахунки для кожного з досліджуваних підходів окремо.

а) Оцінка для ЄС (AI Act)

Спершу відповідно індивідуально для ЄС (AI Act) було виставлено оцінки від п’яти експертів. А після проведено відповідні розрахунки та отримано значення вагових коефіцієнтів.

Таблиця 3.28 – Значення вагових коефіцієнтів для ЄС (AI Act)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	10	6	9	8	7	49	0,1836	0,20408	0,12248	0,1836	0,1632	0,1428
2	10	9	5	10	8	6	48	0,20833	0,1875	0,10416	0,20833	0,1667	0,125
3	8	10	7	9	8	8	50	0,16	0,2	0,14	0,18	0,16	0,16
4	9	10	6	9	8	7	49	0,1836	0,204	0,1224	0,1836	0,16326	0,14285
5	10	10	6	10	9	7	52	0,1923	0,1923	0,11538	0,1923	0,17307	0,1346
							$\sum_{i=1}^{n} r_j$	0,92798	0,98797	0,6044	0,9479	0,8262	0,7053
							w_j	0,18559	0,19759	0,1208	0,1895	0,1652	0,14106

Після отримання усіх результатів було проведено оцінювання – рис. 3.6.

w3 := (0.16444444 0.18111111 0.12666667 0.21444444 0.1544444 0.15388889)

$$L\_Point := \begin{pmatrix} 0.18559759 & 0.19759419 & 0.12088985 & 0.1895959 & 0.16525484 & 0.141006593 \\ 0.18368019 & 0.18309992 & 0.16125155 & 0.17913884 & 0.13563563 & 0.15719386 \\ 0.19817158 & 0.16730776 & 0.18492637 & 0.14574157 & 0.14080027 & 0.16305244 \\ 0.20007323 & 0.1744411 & 0.14047564 & 0.19998073 & 0.13187249 & 0.1531568 \\ 0.16796373 & 0.17222652 & 0.20166662 & 0.13462316 & 0.1301903 & 0.19332967 \\ 0.16244218 & 0.15809436 & 0.19657724 & 0.12821462 & 0.20508788 & 0.14958372 \\ 0.13253901 & 0.15516997 & 0.18277929 & 0.10947515 & 0.21906104 & 0.20097554 \end{pmatrix}$$

V_Point := w3

V_Point = (0.164 0.181 0.127 0.214 0.154 0.154)

Point_rez := L_Point·V_Point^T

Point_rez^T = (0.169 0.167 0.164 0.169 0.163 0.162 0.161)

Рисунок 3.6 – Обчислення результуючого вектору пріоритетів

Найкращий результат за розрахунками отримали ЄС (AI Act) та Велика Британія (White Paper) – 0,169. Найгірший результат показав китайський метод – 0,161. Хоча варто помітити, що розходження між усіма методами – мінімальні: США (NIST AI RMF) – 0,167; ISO/IEC 23894 – 0,164; Канада (AIDA) – 0,163; Південна Корея (AI Basic Act) – 0,162.

Оскільки за результатами хорошу оцінку отримав також підхід Великої Британії, нижче наведено розрахунки і для нього – табл. 3.29.

Таблиця 3.29 – Значення вагових коефіцієнтів Велика Британія (White Paper)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	8	8	10	7	6	9	48	0,1666	0,1666	0,2083	0,1458	0,125	0,1875
2	7	8	9	7	6	9	46	0,1521	0,1739	0,19565	0,1521	0,1304	0,1956
3	8	9	10	6	7	9	49	0,1632	0,1836	0,204	0,1224	0,1428	0,1836
4	9	8	9	6	6	10	48	0,1875	0,1666	0,1875	0,125	0,125	0,2083
5	8	8	10	6	6	9	47	0,17021	0,1702	0,2127	0,1276	0,1276	0,1914
							$\sum_{i=1}^{n} r_j$	0,8398	0,861	1,0083	0,6731	0,6509	0,9666
							w_j	0,16796	0,1722	0,2016	0,1346	0,1301	0,1933

Китайський підхід показав останній результат, визначення його вагових коефіцієнтів та відповідно оцінки експертів наведені у табл. 3.30. Решта розрахунків для кожного з підходів наведено у додатку А.

Таблиця 3.30 – Значення вагових коефіцієнтів (китайський підхід)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	6	7	8	5	10	9	45	0,1333	0,1555	0,1777	0,1111	0,2222	0,2
2	5	7	8	5	10	9	44	0,1136	0,15909	0,1818	0,1136	0,2272	0,2045
3	6	6	8	4	9	9	42	0,1428	0,1428	0,1904	0,0952	0,2142	0,2142
4	6	7	8	5	9	8	43	0,1395	0,1627	0,18604	0,1162	0,2093	0,18604
5	6	7	8	5	10	9	45	0,1333	0,1555	0,1777	0,1111	0,2222	0,2
							$\sum_{i=1}^{n} r_j$	0,6626	0,7758	0,91389	0,5473	1,0953	1,00487
							w_j	0,1325	0,1551	0,1827	0,1094	0,21906	0,2009

3.4.8 Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі числового способу

Останнім застосованим методом став метод визначення вагових коефіцієнтів на основі числового способу. У цьому методі для кожного показника обчислюється коефіцієнт відносного розкиду за формулою [27, 28]:

$$\delta_i = \frac{x_{i\max} - x_{i\min}}{x_{i\max}}.$$

Оскільки значення самих показників можна було отримати з будь-якого попереднього обчислення, були обрані значення обчислень за шкалою Фішберна. А після проведення обчислення вагових коефіцієнтів за формулою [27, 28]:

$$w_i = \frac{\delta_i}{\sum_{i=1}^{m} \delta_i}.$$

Розрахунки вагових коефіцієнтів для критеріїв наведено у табл. 3.31.

Таблиця 3.31 – Значення вагових коефіцієнтів

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,0952381	0,23809524	0,0952381	0,1428571	0,0476191	0,0476191
$x_{i\max}$	0,23809524	0,28571429	0,1904762	0,2857143	0,2380952	0,1904762
δ_i	0,59999998	0,16666667	0,5	0,5	0,8	0,75
w_i	0,18090452	0,05025126	0,1507538	0,1507538	0,241206	0,2261307

Відповідно до результатів оцінювання методом визначення вагових коефіцієнтів на основі числового способу найкращий результат показав ЄС (AI Act) та китайський підхід, розрахунки для обох підходів наведено у табл. 3.32-3.33. Решта розрахунків для кожного з підходів наведено у додатку А.

Таблиця 3.32 – Значення вагових коефіцієнтів для ЄС (AI Act)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,19047619	0,19047619	0,0952381	0,1428571	0,0476191	0,0476191
$x_{i\max}$	0,28571429	0,28571429	0,1904762	0,2857143	0,1428571	0,1428571
δ_i	0,33333334	0,33333335	0,5	0,5	0,6666666	0,6666666
w_i	0,11111111	0,11111112	0,1666667	0,1666667	0,2222222	0,2222222

Таблиця 3.33 – Значення вагових коефіцієнтів китайського підходу

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,19047619	0,0952381	0,0476191	0,1904762	0,1904762	0,0476191
$x_{i\max}$	0,28571429	0,14285714	0,0952381	0,2857143	0,2380952	0,0952381
δ_i	0,33333334	0,33333329	0,5	0,3333333	0,2	0,5
w_i	0,15151515	0,39215679	0,5882353	0,3921569	0,2352941	0,5882353

w4 := (0.18090452 0.05025126 0.1507538 0.1507528 0.241206 0.2261307)

$$L\_Num := \begin{pmatrix} 0.11111111 & 0.11111112 & 0.1666667 & 0.1666667 & 0.2222222 & 0.2222222 \\ 0.16901409 & 0.17605634 & 0.1690141 & 0.1690141 & 0.1760563 & 0.1408451 \\ 0.06944445 & 0.16666668 & 0.2777778 & 0.1388889 & 0.2083333 & 0.1388889 \\ 0.08450704 & 0.08450705 & 0.084507 & 0.1056338 & 0 & 0 \\ 0.125 & 0.08450705 & 0.1056338 & 0.1056338 & 0 & 0 \\ 0.23529411 & 0.47058824 & 0.2941177 & 0 & 0 & 0 \\ 0.15151515 & 0.39215679 & 0.5882353 & 0.3921569 & 0.2352941 & 0.5882353 \end{pmatrix}$$

V_Num := w4

V_Num = (0.181 0.05 0.151 0.151 0.241 0.226)

$Num\_rez := L\_Num \cdot V\_Num^T$

$Num\_rez^T = (0.18 \;\; 0.165 \;\; 0.165 \;\; 0.048 \;\; 0.059 \;\; 0.111 \;\; 0.385)$

Рисунок 3.7 – Обчислення результуючого вектору пріоритетів

Найкращий результат за розрахунками отримали китайський підхід – 0,385 та ЄС (AI Act) – 0,18. Найгірший результат показав Велика Британія (White Paper) – 0,048. Інші методи отримали наступні результати: США (NIST AI RMF) – 0,156; ISO/IEC 23894 – 0,156; Канада (AIDA) – 0,059; Південна Корея (AI Basic Act) – 0,111.

3.4.9 Висновки за результатами оцінювання

Провівши оцінювання адекватності різних міжнародних підходів регулювання штучного інтелекту щодо застосування в Україні ми отримали результати, які схематично наведені на рис. 3.8.

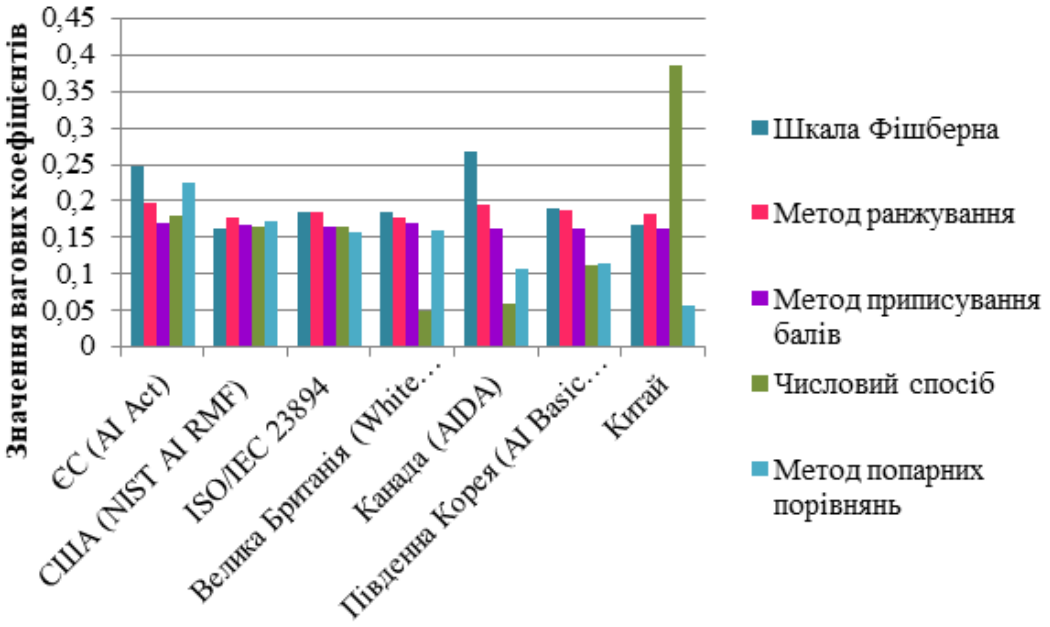

Рисунок 3.8 – Ілюстрація результатів оцінювання

Оцінювання проводилось за наступними критеріями, які ми вважаємо важливими при теоретичному огляді можливості інтеграції: рівень прозорості алгоритмів і систем прийняття рішень, тобто пояснюваність процедур, вимог та інших дій зазначених у документі; перспективність інтеграції регуляторного акту, методу, стандарту в Україні; рівень гнучкості документу у застосуванні в різних сферах; можливість запобігання виникненню етичних ризиків, пов'язаних з порушенням прав та свобод людини; рівень інституційної підтримки та державної залученості; вплив на перспективність розвитку технології окремо та у інших технологічних галузях.

За результатами описаними раніше та наведеними на рис. 3.8, ми бачимо, що найкраще за різними методами оцінювання виявились ЄС (AI Act) та Канада (AIDA). При чому ЄС (AI Act) показує доволі стабільні значення, що свідчить про збалансованість та зрілість регуляторної системи, де враховуються як прозорість алгоритмів, так і етична складова та інституційна підтримка.

Китайський підхід до регулювання штучного інтелекту показав найгірші результати майже за всіма методиками. Тож є одним із найменш ймовірно інтегрованих для України. Хоча за аналізом усереднених значень, наведеному у табл. 3.34 та рис. 3.9, цей підхід показав другий результат.

США (NIST AI RMF) та ISO/IEC 23894 мають близькі результати, з незначною перевагою першого. Обидва документи роблять акцент на гнучкості й стандартизації. Це робить їх доволі хорошими прикладами для України.

Велика Британія демонструє середні показники за всіма методами, що відповідає її гнучкому, ринково-орієнтованому підходу, заснованому на принципах «регулювання через довіру». Цей підхід забезпечує помірну ефективність, але нижчу централізованість.

Південна Корея має дещо нижчі значення, однак показує відносну збалансованість, але все ж це робить її не найкращим прикладом для наслідування у випадку України.

Усереднене оцінювання результатів для кожного з підходів наведено у табл. 3.34.

Таблиця 3.34 – Усереднене оцінювання результатів

Методи (документи)	Середній результат
ЄС (AI Act)	0,2038
США (NIST AI RMF)	0,1684
ISO/IEC 23894	0,1716
Велика Британія (White Paper)	0,1474
Канада (AIDA)	0,1584
Південна Корея (AI Basic Act)	0,1532
Китай	0,1908

Ілюстративний аналіз усереднених значень оцінювання наведено на рис. 3.9.

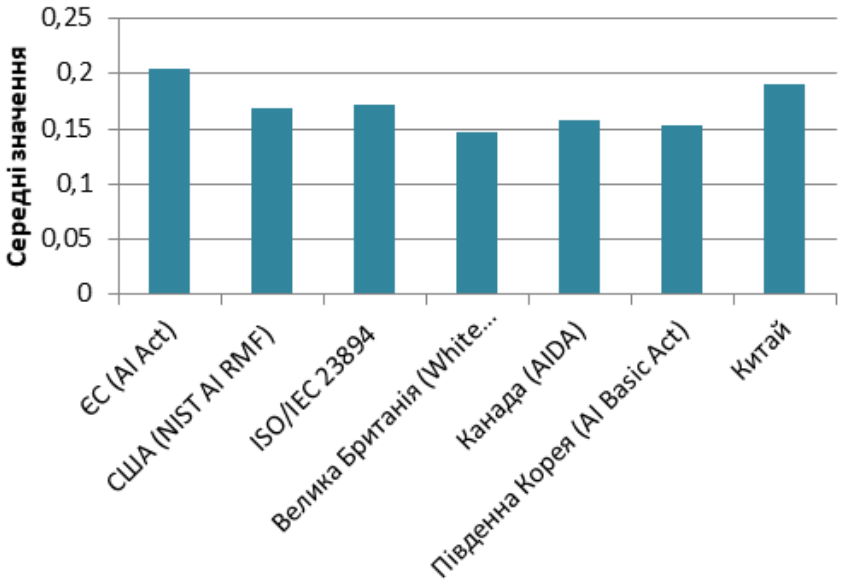

Рисунок 3.9 – Усереднене оцінювання результатів

ВИСНОВКИ

Розвиток технологій штучного інтелекту є не лише потужним рушієм цифрової трансформації, а й серйозним викликом у сфері безпеки. Вивчення міжнародних підходів показало, що жодна країна не має абсолютно універсального методу ефективного регулювання ШІ, однак більшість систем сходяться у спільному прагненні – забезпечити баланс між технологічним прогресом, правами людини та етичною відповідальністю перед суспільством. Саме цей баланс має стати основою формування моделі безпеки штучного інтелекту для України.

Аналіз сучасного стану правового регулювання в Україні показав, що країна перебуває на етапі становлення власної політики у сфері штучного інтелекту. Існує стратегічне бачення інтеграції в європейський цифровий простір, поступового впровадження стандартів AI Act і формування гнучкого правового середовища, яке сприятиме розвитку інновацій. Позитивним кроком є розробка «Білої книги», створення регуляторної пісочниці та ініціатив із добровільного маркування систем. Такі кроки допоможуть сформувати культуру відповідального використання технологій і забезпечити довіру суспільства до цифрових рішень.

Запропонована модель безпеки ШІ для України передбачає системну взаємодію між державними структурами, академічною спільнотою, бізнесом і громадянським суспільством. Її основою є багаторівнева система оцінки ризиків, запровадження етичного реєстру моделей ШІ та створення спеціального державного департаменту з питань штучного інтелекту. Важливими елементами моделі є обов’язкове маркування контенту, прозорість алгоритмів і сертифікація безпеки. Такий підхід дозволяє не лише мінімізувати ризики, але й створити умови для розвитку конкурентного ринку інноваційних рішень.

Проведений порівняльний аналіз засвідчив, що підхід Європейського Союзу (AI Act) є найбільш структурованим і детальним, ґрунтується на принципі ризик-орієнтованості та підзвітності. Він придатний для імплементації в українське

правове поле, адже відповідає курсу євроінтеграції та підходам до захисту прав людини.

Американська модель NIST AI RMF, хоч і є добровільною, демонструє ефективність у сфері корпоративного управління ризиками й може бути використана для розробки внутрішніх політик підприємств.

Британський White Paper підкреслює важливість інноваційності та децентралізації, тоді як канадський AIDA концентрується на прозорості, етичності та відповідальності розробників.

Південна Корея, натомість, вибудувала високорівневу систему державного контролю та підтримки інновацій, а Китай – приклад централізованої, контрольованої моделі, спрямованої на стабільність і безпеку.

Проведений аналіз свідчить, що ефективне регулювання ШІ неможливе без міждисциплінарного підходу. Необхідно поєднувати технічні стандарти з правовими, етичними та соціальними аспектами.

Для України пріоритетом має стати формування власних галузевих стандартів безпеки, гармонізованих із міжнародними документами, такими як ISO/IEC 23894, AI Act, AIDA. Впровадження цих принципів дозволить забезпечити сумісність українських рішень із глобальними екосистемами та сприятиме залученню іноземних інвестицій у технологічний сектор.

Таким чином, результати дослідження підтверджують, що створення комплексної системи безпеки штучного інтелекту є стратегічно важливим завданням для України. Вона повинна поєднувати ризик-орієнтоване регулювання, етичні стандарти, інституційну підтримку та розвиток інновацій. Реалізація запропонованих рекомендацій сприятиме не лише зміцненню національної кібербезпеки, але й формуванню довіри до цифрової держави.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Cost of a Data Breach Report 2025. [Електронний ресурс]. – Режим доступу: https://www.ibm.com/reports/data-breach.

2. What is the history of artificial intelligence (AI)? [Електронний ресурс]. – Режим доступу: https://www.tableau.com/data-insights/ai/history.

3. Charles Babbage's Difference Engines and the Science Museum [Електронний ресурс]. – Режим доступу: https://www.sciencemuseum.org.uk/objects-and-stories/charles-babbages-difference-engines-and-science-museum.

4. Understanding the different types of artificial intelligence. [Електронний ресурс]. – Режим доступу: https://www.ibm.com/think/topics/artificial-intelligence-types.

5. What is reasoning in AI? [Електронний ресурс]. – Режим доступу: https://www.ibm.com/think/topics/ai-reasoning.

6. Логачова Є. О. Дослідження та аналіз міжнародних стандартів та регуляторних вимог щодо безпеки штучного інтелекту, розробка моделі безпеки для України / Є. О. Логачова, М. В. Єсіна, Д. Ю. Голубничий // Радіотехніка. – Х. : Харківський національний університет радіоелектроніки, 2025. – Випуск 221. – С. 14-22.

7. Логачова Є. О. Штучний інтелект: сучасні загрози та безпека / Є. О. Логачова, М. В. Єсіна // X Міжнародна науково-технічна конференція «Комп'ютерне моделювання в наукоємних технологіях»: Праці X міжнародної науково-технічної конференції, 27-29 листопада 2024 року. Харків: ХНУ імені В. Н. Каразіна, 2024. – С. 123-126.

8. EU AI Act: first regulation on artificial intelligence. [Електронний ресурс]. – Режим доступу: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

9. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. – Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

10. Information technology – Artificial intelligence – Guidance on risk management. – Режим доступу: https://cdn.standards.iteh.ai/samples/77304/cb803ee4e9624430a5db177459158b24/ISO-IEC-23894-2023.pdf.

11. Trump unveils AI plan that aims to clamp down on regulations and 'bias'. [Електронний ресурс]. – Режим доступу: https://www.bbc.com/news/articles/c4g8nxrk207o.

12. AI Watch: Global regulatory tracker – United Arab Emirates. [Електронний ресурс]. – Режим доступу: https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-uae.

13. Albania turns to AI to beat corruption and join EU. [Електронний ресурс]. – Режим доступу: https://www.politico.eu/article/albania-use-ai-artificial-intelligenve-join-eu-corruption/.

14. Interim Measures for the Management of Generative Artificial Intelligence Services. [Електронний ресурс]. – Режим доступу: https://www.chinalawtranslate.com/en/generative-ai-interim/#gsc.tab=0.

15. A pro-innovation approach to AI regulation 3 August 2023. [Електронний ресурс]. – Режим доступу: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper?utm_source=chatgpt.com#part-7-conclusion-and-next-steps.

16. The Artificial Intelligence and Data Act (AIDA) – Companion document. [Електронний ресурс]. – Режим доступу: https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document.

17. South Korea Artificial Intelligence (AI) Basic Act. [Електронний ресурс]. – Режим доступу: https://www.trade.gov/market-intelligence/south-korea-artificial-intelligence-ai-basic-act.

18. Логачова Є. О. Інформаційна модель регулювання безпеки штучного інтелекту в Україні / Є. О. Логачова, М. В. Єсіна // ІІІ Міжнародна науково-практична конференція з проведенням CTF-змагання: Праці ІІІ міжнародної науково-технічної конференції, 6-7 травня 2025 року. К. : Військовий інститут телекомунікації та інформатизації імені Героїв Крут, 2024. – С. 33-34.

19. Біла книга з регулювання ШІ в Україні: бачення Мінцифри. Версія для консультацій. – Режим доступу: https://thedigital.gov.ua/storage/uploads/files/page/community/docs/%D0%A0%D0%B5%D0%B3%D1%83%D0%BB%D1%8E%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F%20%D0%A8%D0%86.pdf.

20. Дорожня карта з регулювання штучного інтелекту в Україні. [Електронний ресурс]. – Режим доступу: https://surl.li/cgvecu.

21. Закон України «Про рекламу». [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/270/96-%D0%B2%D1%80#Text.

22. Закон України «Про захист прав споживачів». [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/1023-12#Text.

23. Закон України «Про стандартизацію». [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/1315-18#Text.

24. Закон України «Про охорону прав на винаходи і корисні моделі». [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/3687-12#Text.

25. Цивільний кодекс України. [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/435-15#Text.

26. Стаття 42, Конституція України. [Електронний ресурс]. – Режим доступу: https://constitution.in.ua/articles/42/.

27. Горбенко І. Д. Методи, методика та результати порівняльного аналізу електронних підписів згідно ДСТУ ISO/IEC 14888-3:2014 / І. Д. Горбенко, М. В. Єсіна // Вісник Національного університету “Львівська політехніка”: серія “Автоматика, вимірювання та керування”. – Л. : Національний університет “Львівська політехніка”, 2016. – № 852 – С. 9–22.

28. Горбенко Ю. І., Ганзя Р. С., Акользіна О. С. Електронні підписи на основі ідентифікаторів та бінарного відображення // Прикладна радіоелектроніка. Х. : Харківський національний університет радіоелектроніки, 2015. Т. 14. № 4. С. 284–290.

29. Oletsky O. Enhancing Consistency of Pairwise Comparisons on the Base of Linear Algebraic Equations. NaUKMA Research Papers. Computer Science. 2023. T. 5. C. 85–91.

30. Lyashenko, E. M., Klymash, M. P., Lyashko, O. A. On creation of expert knowledge base for estimation of state of organizational systems // International Journal of Fuzzy Systems and Advanced Applications. – 2014. – Vol. 1. – P. 7–13.

31. Alguliyev R. M., Nabibayova G. Ch., Abdullayeva S. R. Evaluation of websites by many criteria using the algorithm for pairwise comparison of alternatives // I.J. Intelligent Systems and Applications. – 2020. – Vol. 12, No. 6. – P. 64–74. https://doi.org/10.5815/ijisa.2020.06.05.

ДОДАТОК А

РОЗРАХУНКИ ДО РОЗДІЛУ 3

1. Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі методу приписування балів

Таблиця 1.1 – Значення вагових коефіцієнтів США (NIST AI RMF)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	8	8	7	9	6	7	45	0,17777778	0,17777778	0,15555556	0,2	0,13333333	0,15555556
2	9	9	8	7	6	7	46	0,19565217	0,19565217	0,17391304	0,15217391	0,13043478	0,15217391
3	8	9	9	8	6	7	47	0,17021277	0,19148936	0,19148936	0,17021277	0,12765957	0,14893617
4	9	7	6	8	7	7	44	0,20454545	0,15909091	0,13636364	0,18181818	0,15909091	0,15909091
5	8	9	7	9	6	8	47	0,17021277	0,19148936	0,14893617	0,19148936	0,12765957	0,17021277
							$\sum_{i=1}^{n} r_j$	0,91840094	0,91549958	0,80625777	0,89569422	0,67817817	0,78596931
							w_j	0,18368019	0,18309992	0,16125155	0,17913884	0,13563563	0,15719386

Таблиця 1.2 – Значення вагових коефіцієнтів ISO/IEC 23894

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	8	9	7	7	8	48	0,1875	0,16666667	0,1875	0,14583333	0,14583333	0,16666667
2	10	8	9	6	7	7	47	0,21276596	0,17021277	0,19148936	0,12765957	0,14893617	0,14893617
3	8	7	8	7	6	7	43	0,18604651	0,1627907	0,18604651	0,1627907	0,13953488	0,1627907
4	9	7	8	7	6	7	44	0,20454545	0,15909091	0,18181818	0,15909091	0,13636364	0,15909091
5	9	8	8	6	6	8	45	0,2	0,17777778	0,17777778	0,13333333	0,13333333	0,17777778
							$\sum_{i=1}^{n} r_j$	0,99085792	0,83653882	0,92463183	0,72870785	0,70400136	0,81526222
							w_j	0,19817158	0,16730776	0,18492637	0,14574157	0,14080027	0,16305244

Таблиця 1.3 – Значення вагових коефіцієнтів Канада (AIDA)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	8	7	10	6	7	47	0,19148936	0,17021277	0,14893617	0,21276596	0,12765957	0,14893617

2	10	8	6	9	6	7	46	0,2173913	0,17391304	0,13043478	0,19565217	0,13043478	0,15217391
3	9	8	7	10	6	8	48	0,1875	0,16666667	0,14583333	0,20833333	0,125	0,16666667
4	9	8	7	9	6	7	46	0,19565217	0,17391304	0,15217391	0,19565217	0,13043478	0,15217391
5	10	9	6	9	7	7	48	0,20833333	0,1875	0,125	0,1875	0,14583333	0,14583333
							$\sum_{i=1}^{n} r_j$	1,00036617	0,87220552	0,7023782	0,99990364	0,65936247	0,765784
							w_j	0,20007323	0,1744411	0,14047564	0,19998073	0,13187249	0,1531568

Таблиця 1.4 – Значення вагових коефіцієнтів Південна Корея (AI Basic Act)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	8	7	9	6	9	7	46	0,17391304	0,15217391	0,19565217	0,13043478	0,19565217	0,15217391
2	7	8	9	6	10	7	47	0,14893617	0,17021277	0,19148936	0,12765957	0,21276596	0,14893617
3	8	7	9	6	10	7	47	0,17021277	0,14893617	0,19148936	0,12765957	0,21276596	0,14893617
4	7	8	10	6	9	7	47	0,14893617	0,17021277	0,21276596	0,12765957	0,19148936	0,14893617
5	8	7	9	6	10	7	47	0,17021277	0,14893617	0,19148936	0,12765957	0,21276596	0,14893617
							$\sum_{i=1}^{n} r_j$	0,81221092	0,79047179	0,98288622	0,64107308	1,02543941	0,74791859
							w_j	0,16244218	0,15809436	0,19657724	0,12821462	0,20508788	0,14958372

2. Оцінювання адекватності застосування міжнародних підходів методом визначення вагових коефіцієнтів на основі числового способу

Таблиця 2.1 – Значення вагових коефіцієнтів для США (NIST AI RMF)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,04761905	0,04761905	0,0476191	0,0476191	0,0476191	0,0952381
$x_{i\max}$	0,23809524	0,28571429	0,2380952	0,2380952	0,2857143	0,2857143
δ_i	0,8	0,83333333	0,8	0,8	0,8333333	0,6666667
w_i	0,16901409	0,17605634	0,1690141	0,1690141	0,1760563	0,1408451

Таблиця 2.2 – Значення вагових коефіцієнтів для ISO/IEC 23894

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,23809524	0,14285714	0,0476191	0,1904762	0,0476191	0,0952381
$x_{i\max}$	0,28571429	0,23809524	0,1428571	0,2857143	0,0952381	0,1428571
δ_i	0,16666667	0,40000002	0,6666666	0,3333333	0,5	0,3333333
w_i	0,06944445	0,16666668	0,2777778	0,1388889	0,2083333	0,1388889

Таблиця 2.3 – Значення вагових коефіцієнтів для Великої Британії (White Paper)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,14285714	0,14285714	0,1428571	0,1428571	0,0952381	0,0476191
$x_{i\max}$	0,23809524	0,23809524	0,2380952	0,2857143	0,0952381	0,0476191
δ_i	0,4	0,40000002	0,4	0,5	0	0
w_i	0,08450704	0,08450705	0,084507	0,1056338	0	0

Таблиця 2.4 – Значення вагових коефіцієнтів для Канади (AIDA)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,19047619	0,14285714	0,1428571	0,1428571	0,0952381	0,0476191
$x_{i\max}$	0,23809524	0,23809524	0,2857143	0,2857143	0,0952381	0,0476191
δ_i	0,2	0,40000002	0,5	0,5	0	0
w_i	0,125	0,08450705	0,1056338	0,1056338	0	0

Таблиця 2.5 – Значення вагових коефіцієнтів для Південної Кореї (AI Basic Act)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,19047619	0,14285714	0,1428571	0,2857143	0,0476191	0,0952381
$x_{i\max}$	0,23809524	0,23809524	0,1904762	0,2857143	0,0476191	0,0952381
δ_i	0,2	0,40000002	0,25	0	0	0
w_i	0,23529411	0,47058824	0,2941177	0	0	0

ДОДАТОК Б

ПЕРЕЛІК ПУБЛІКАЦІЙ

МІНІСТЕРСТВО ОБОРОНИ УКРАЇНИ

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут

III МІЖНАРОДНА НАУКОВО-ПРАКТИЧНА КОНФЕРЕНЦІЯ

The third International scientific and practical conference

“Кіберборотьба: розвідка, захист та протидія”

“Cyberwarfare: intelligence, defense and offensive security”

06–07.05.2025

ТЕЗИ ДОПОВІДЕЙ

Київ

Логачова Є. О. (ХНУ імені В. Н. Каразіна)
канд. техн. наук, доцент **Єсіна М. В.** (ХНУ імені В. Н. Каразіна)

ІНФОРМАЦІЙНА МОДЕЛЬ РЕГУЛЮВАННЯ БЕЗПЕКИ ШТУЧНОГО ІНТЕЛЕКТУ В УКРАЇНІ

Вступ. Розумні технології, що ґрунтуються на штучному інтелекті (ШІ), стрімко поширюються в різних сферах життя – від повсякденного використання до медицини та сфери безпеки. Водночас зростає потреба в забезпеченні належного рівня безпеки цих технологій. У відповідь на це в Європейському Союзі та низці інших країн вже з'явилися перші нормативно-правові акти, спрямовані на регулювання ШІ. В Україні ж поки що відсутнє єдине бачення щодо необхідності запровадження подібних регулювань, хоча перші кроки у цьому напрямі вже зроблено.

1. Класифікація ризиків, пов'язаних із використанням штучного інтелекту

Перші спроби створення ШІ були ще у 90-х роках минулого століття, та найбільшим моментом прогресу вважається 2023 рік, у якому відбувся випуск GPT-3 та його наступників. А вже у 2024 році ці розробки показують неабиякий прогрес у генеруванні зображень, текстів, аудіо й іншого. До того ж було помічено прогрес у навчанні ШІ, завдяки якому помічники на основі ШІ здатні брати участь у природних, контекстних розмовах, а також допомагати з різноманітними завданнями [1].

ШІ несе за собою не менший список загроз і ризиків, наприклад: проблеми з конфіденційністю, алгоритмічна упередженість, неетичне використання, можливі втрати робочих місць, автоматизація кібератак, проблеми довіри та контролю, використання ШІ для атак соціальної інженерії, атаки на самі алгоритми штучного інтелекту та багато іншого. У документі EU AI Act проводиться поділ ШІ за рівнями ризиків, відповідно до яких мають бути задіяні різні правила. За цим розподілом ризики, які несуть за собою технології ШІ, поділяють на три типи: неприйнятний, високий і мінімальний [2].

2. Міжнародні регуляторні заходи з безпеки використання штучного інтелекту

Одним із перших органів, що висловив занепокоєння щодо потенційних ризиків, пов'язаних із ШІ, став Конгрес США. У 2023 році генеральний директор компанії OpenAI Сем Альтман закликав Конгрес запровадити регулювання в цій сфері. Яскравим прикладом таких зусиль стало ухвалення 17 травня 2024 року в штаті Колорадо першого в США комплексного закону про ШІ – Колорадського закону про штучний інтелект. Цей нормативний акт визначає обов'язки як для розробників, так і для користувачів систем ШІ. Основна увага приділяється автоматизованим системам ухвалення рішень. Зокрема, під поняття «високоризикована система ШІ» підпадають ті, що під час використання ухвалюють рішення або суттєво впливають на остаточне рішення. Закон набуде чинності у 2026 році [3]. Ще одним стандартом у цій галузі є NIST AI Risk Management Framework [4]. Це методичний посібник, створений для організацій, які використовують або розробляють генеративний ШІ (GAI). Він допомагає: визначити та зрозуміти ризики, пов'язані з GAI (наприклад, deepfakes, порушення приватності, токсичний контент тощо); розробити політики, стратегії та дії для ефективного управління цими ризиками; дотримуватись міжнародних стандартів і принципів довіри, безпеки та прозорості ШІ [4].

Регулювання, введені у ЄС, мають дещо інший характер. У країнах ЄС закликають вводити певні регулювання і обмеження не тільки на рівні держав у контексті ШІ, а й на приватному громадському рівні. Регламент ЄС 2024/1689 є важливим правовим документом Європейського Союзу, який встановлює гармонізовані правила для використання ШІ у різних сферах. Його основною метою є створення єдиного правового поля для регулювання ШІ у країнах-членах ЄС. Регламент зобов'язує розробників і постачальників ШІ забезпечувати можливість людського контролю, мінімізувати дискримінаційні ризики та захищати права людини. Документ також вносить зміни до кількох існуючих європейських регламентів, таких як транспорт, захист даних і безпека продукції, щоб узгодити вимоги ШІ з чинним законодавством. Наступним важливим рішенням стала директива NIS2, яка не фокусується

безпосередньо на регулюванні ШІ, але одним із ключових аспектів є те, що NIS2 розширює вимоги до кібербезпеки для критичних інфраструктур і послуг, включаючи ті, де використовуються технології ШІ [5].

3. Інформаційна модель регулювання безпеки штучного інтелекту в Україні

Попри відсутність активної публічної дискусії щодо регулювання ШІ в Україні, певні ініціативи в цьому напрямку вже реалізуються. Зокрема, Україна приєдналася до Декларації Bletchley, підписаної під час Саміту з безпеки ШІ у 2023 році. Крім того, була розроблена так звана Біла книга, яка має на меті сприяти конкурентоспроможності українського бізнесу, забезпечити захист прав людини та підтримати процес євроінтеграції. У документі пропонується етапний підхід: від використання позазаконодавчих інструментів та ініціатив у найближчі роки – до ухвалення спеціального закону про ШІ на завершальному етапі. Водночас запропонована модель прагне вибудувати систему регулювання безпеки ШІ в Україні більш узгоджено та інтегровано із міжнародними практиками [5].

У моделі пропонується додати більше суб'єктів діяльності, пов'язаної з ШІ, а також інтегрувати вже відомий з EU AI Act розподіл за рівнями ризику. Запропоновану модель наведено на рисунку 1.

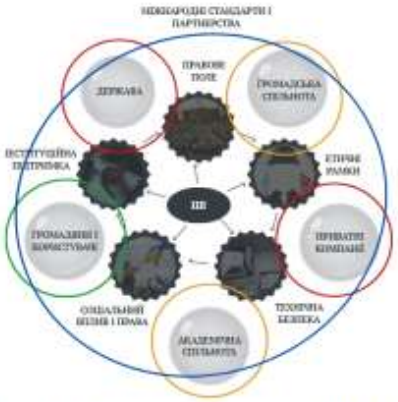

Рис. 1. Модель регулювання безпеки ШІ в Україні

Висновки. Швидкий розвиток технологій ШІ вимагає від держав комплексного підходу до безпеки і регулювання, з урахуванням міжнародного досвіду та національних реалій. Україна, попри відсутність усталеного законодавчого поля у сфері ШІ, уже робить кроки у напрямку гармонізації з європейськими стандартами. Запропонована інформаційна модель регулювання дозволяє систематизувати ризики, окреслити ролі ключових суб'єктів і сформувати адаптивну систему, орієнтовану на захист прав людини, кібербезпеку та інноваційний розвиток.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. What is artificial intelligence (AI)? // ISO. 2024. URL: https://www.iso.org/artificial-intelligence/what-is-ai-v3.

2. EU AI Act: first regulation on artificial intelligence. European Parliament, 2025. URL: https://surl.li/abxfru.

3. AI Watch: Global regulatory tracker – United States. White & Case LLP, 2024. URL: https://surl.li/snkzfd.

4. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, U.S. Department of Commerce Gina M. Raimondo, Secretary NIST Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology, 2024. URL: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

5. Логачова Є. О, Єсіна М. В. Штучний інтелект: сучасні загрози та безпека. 2024.

УДК 004.056.5:004.8 DOI:10.30837/rt.2025.2.221.02

Є.О. ЛОГАЧОВА, М.В. ЄСІНА, канд. техн. наук, Д.Ю. ГОЛУБНИЧИЙ, канд. техн. наук

ДОСЛІДЖЕННЯ ТА АНАЛІЗ МІЖНАРОДНИХ СТАНДАРТІВ ТА РЕГУЛЯТОРНИХ ВИМОГ ЩОДО БЕЗПЕКИ ШТУЧНОГО ІНТЕЛЕКТУ, РОЗРОБКА МОДЕЛІ БЕЗПЕКИ ДЛЯ УКРАЇНИ

Вступ

В епоху сучасних технологій надважливо забезпечити їх необхідним рівнем безпеки для збереження належного функціонування бізнесу, різноманітних структур тощо. Також необхідно запобігти порушенню прав та свобод звичайних громадян і користувачів. Штучний інтелект (ШІ) – відносно нова технологія на ринку, що завоювала прихильність як звичайних користувачів, так і великих корпорацій. Розумні машини на основі штучного інтелекту допомогли у стрімкому розвитку багатьох галузей, таких як медицина, освіта, транспорт, сільське господарство, фінанси тощо. Разом з тим ШІ почали використовувати і для покращення рівня кібершахрайств. Для того щоб не відмовлятись від даних прогресивних технологій та не уповільнювати прогрес, багато країн вже почали вводити різноманітні рішення щодо безпеки використання штучного інтелекту. Україна наразі знаходиться на початковій стадії у даному питання, тому розгляд міжнародного досвіду допоможе зробити нормативно-правові регулювання в Україні ефективними та безпечними для усіх ланок держави.

1. Кібербезпека в епоху штучного інтелекту

Штучний інтелект радикально трансформує цифровий світ: автоматизація процесів, аналіз великих даних, прогнозування загроз. Спершу ШІ використовували для маленьких повторюваних задач, на зараз штучний інтелект використовується для діагностики здоров'я, організації бізнесу, навчання. Тобто технологія розвивається стрімко і використовується майже у всіх сферах життя, чи неабияк полегшує життя своїх користувачів.

За даними IBM, середня вартість витоку даних у 2023 р. становила понад 4,45 млн доларів, і ШІ дедалі частіше використовується як захисний, так і атакуючий інструмент [1]. А вже у 2024 р., згідно з IBM Cost of a Data Breach Report 2024, середня вартість витоку даних зросла до 4,88 млн доларів США, що на 10 % більше порівняно з 2023 р. Та при цьому штучний інтелект та автоматизація допомогли компаніям знизити витрати на усунення наслідків витоку даних у середньому на 2,2 млн доларів [2]. ШІ допомагає вирішувати декілька основних проблем інформаційної безпеки: виявлення загроз у реальному часі, автоматичне реагування на інциденти, прогнозування атак та інші.

Але з розвитком ШІ зростають і виклики – зокрема в сфері кібербезпеки. Відтак зловмисники можуть використати технології штучного інтелекту для створення deepfake контенту, автоматизації фішингових атак, атак соціальної інженерія нового рівня тощо. Окрім цього дана технологія може бути задіяна і у автоматизації шкідливого програмного забезпечення (ПЗ). Один із показових прикладів – дослідницький експеримент фахівців з кібербезпеки компанії Hyas, які створили умовний вірус під назвою BlackMamba. Цей шкідливий код мав здатність динамічно змінювати свою поведінку за допомогою ChatGPT, генеруючи частини свого функціоналу в режимі реального часу. Під час випробувань BlackMamba демонстрував здатність адаптуватися до різних середовищ та залишатися непоміченим більшістю популярних антивірусних систем [3].

У лютому 2024 р. в Гонконзі стався один із наймасштабніших випадків шахрайства із використанням штучного інтелекту. Співробітник міжнародної компанії отримав завдання надіслати 25 млн доларів США від імені свого нібито директора з Великої Британії. Після сумнівного електронного листа із запрошенням на відеоконференцію чоловік почав підозрювати фішинг. Проте участь кількох «знайомих колег» у відеодзвінку знизила його настороженість. Протягом кількох транзакцій, що охоплювали п'ять різних банківських рахунків,

працівник переказав загалом понад 200 млн гонконзьких доларів (приблизно 25 млн). Як згодом з'ясувалося, усі інші учасники відеоконференції були глибоко реалістичними штучними двійниками, створеними зловмисниками з використанням технологій штучного інтелекту та deepfake [4].

Усе це ставить непросту задачу перед спеціалістами з інформаційної безпеки, адже дана технологія неабияк підвищує рівень безпеки та при цьому ж може бути використана кіберзлочинцями для здійснення атак. Хорошим рішенням є впровадження необхідних стандартів та регуляторних вимог для безпечного використання ШІ. Для України це питання стоїть доволі серйозно, адже в умовах війни кількість фейків та кібератак все збільшується.

2. Огляд міжнародних стандартів та регуляторних вимог безпеки ШІ

Існує багато підходів і думок щодо нормативно-правових регулювань штучного інтелекту. У одних країнах до цієї технології ставляться з обережністю, а у інших вбачають у ній великий прогресивний поштовх для розвитку цифровізації.

Варто почати з Європейського Союзу (ЄС) та введеного ними EU AI Act, який фактично став першим офіційним і повноцінним документом з регулювання ШІ. EU AI Act встановлює чіткі правила для використання штучного інтелекту, враховуючи рівень потенційного ризику, який можуть нести різні типи систем ШІ. Залежно від цього рівня, як провайдери, так і користувачі зобов'язані дотримуватися певних вимог. Навіть у випадках, якщо ризик від технології вважається низьким, вона має пройти оцінювання, щоб гарантувати прозорість та відповідність етичним стандартам [5].

До категорії неприйнятного ризику потрапляють застосунки ШІ, які заборонені в ЄС, оскільки становлять загрозу для базових прав і свобод людини. Це включає штучний інтелект, що маніпулює поведінкою людей – особливо вразливих груп, таких як діти. Наприклад, голосові іграшки, які можуть спонукати до небезпечної поведінки. Також заборонені системи соціальної оцінки (оцінювання людей за поведінкою чи статусом), біометрична ідентифікація в публічних просторах у реальному часі, а також категоризація людей за біометричними ознаками [5].

У деяких випадках, як-от діяльність правоохоронних органів, можливе обмежене використання систем з високим ризиком. Наприклад, дистанційна біометрична ідентифікація в реальному часі дозволяється лише у виняткових ситуаціях та після попереднього дозволу суду. Ідентифікація «постфактум», тобто із затримкою, може застосовуватися для розслідування тяжких злочинів, але також лише за наявності юридичного схвалення [5].

Системи штучного інтелекту, які становлять високий ризик, охоплюють сфери, що мають прямий вплив на безпеку чи фундаментальні права громадян. Вони поділяються на дві ключові категорії. Перша – це ШІ, інтегрований у продукти, що регулюються європейським законодавством щодо безпеки. Наприклад, медичне обладнання, автомобілі, іграшки, ліфти. Друга – системи, які функціонують у чутливих галузях: управління критичною інфраструктурою, освіта, зайнятість, доступ до державних послуг, правоохоронна діяльність, а також процеси міграції й правозастосування [5].

Всі ці системи підлягають обов'язковому контролю як до їхнього впровадження на ринку, так і протягом усього життєвого циклу. Крім того, громадяни матимуть право подавати скарги на ШІ-системи до національних наглядових органів, що зміцнює принцип прозорості та підзвітності в епоху технологічного прогресу.

Ще одним важливим принципом є прозорість. Так, наприклад ChatGPT не вважається технологією з високим рівнем ризику, проте має відповідати законодавству ЄС, зокрема законодавству щодо авторського права. Таким чином, увесь контент, згенерований ChatGPT, має бути маркований спеціальним знаком, це допомагає запобігати поширенню фейкових зображень, відео та іншого контенту з шахрайською метою. Більш досконала модель штучного інтелекту GPT-4 повинна проходити ретельну оцінку, а про будь-які серйозні інциденти потрібно повідомляти Європейську комісію [5].

чити надійність, безпечність і етичність ШІ-систем протягом усього їхнього життєвого циклу – від початкового проєктування до постійного оновлення.

Перший етап – Govern, він полягає у створенні чіткої організаційної структури та політик, що регулюють впровадження ШІ. Це включає визначення відповідальних осіб, розподіл повноважень, прозорість процесів і формування внутрішніх стандартів. Добре організоване управління дозволяє приймати обґрунтовані рішення, враховуючи етичні, правові та соціальні аспекти використання технологій [6].

Далі йде Map – функція, що допомагає оцінити контекст, у якому працює система. На цьому етапі організації ідентифікують можливі загрози, враховують потреби користувачів і визначають потенційний вплив ШІ на людей та навколишнє середовище. Важливо не лише знати, як працює система, а й розуміти, для кого і з якою метою вона створена [6].

Measure забезпечує вимірювання ризиків та надійності системи [6]. Це можуть бути кількісні та якісні показники. Наприклад, точність, стабільність, справедливість або захищеність моделі. Регулярна оцінка допомагає виявляти слабкі місця до того, як вони спричинять шкоду, і дає змогу приймати обґрунтовані рішення щодо подальших кроків.

Функція Manage передбачає реалізацію конкретних заходів для мінімізації виявлених ризиків [6]. Це може включати зміни в алгоритмах, обмеження доступу до деяких функцій, захист персональних даних або навіть перегляд стратегії впровадження. Важливо, що управління ризиками не є разовим завданням, а триває протягом усього періоду використання ШІ.

Завершує цей цикл функція Improve, яка зосереджена на постійному аналізі та вдосконаленні практик управління ризиками. Організації мають враховувати досвід, нові знання, технологічний розвиток та зміни в нормативному середовищі, щоб адаптувати свої ШІ-системи до нових викликів [6].

Ще одним документом є стандарт ISO/IEC 23894:2023. Це перший міжнародний стандарт, що надає рекомендації з управління ризиками, пов'язаними з використанням штучного інтелекту. Документ адаптує загальні принципи управління ризиками до специфіки ШІ-систем, враховуючи їхню складність, динамічну поведінку, етичні виклики та вплив на права людини. Головна мета стандарту – допомогти організаціям мінімізувати можливу шкоду, підвищити довіру до ШІ та зробити його застосування безпечнішим і більш передбачуваним [7].

У стандарті ISO/IEC 23894:2023 окреслюється структура ефективного управління ризиками штучного інтелекту, що охоплює весь життєвий цикл систем ШІ. Одним із перших і ключових етапів є ідентифікація ризиків, яка передбачає глибоке розуміння того, як система ШІ буде використовуватися на практиці. Тут організації повинні враховувати як передбачувані сценарії застосування, так і потенційні випадки неправильного або зловмисного використання. Важливо також ретельно проаналізувати дані, на яких навчалась модель, способи прийняття нею рішень та ймовірний вплив її функціонування на різні групи користувачів і суспільство загалом [7].

Після виявлення потенційних ризиків стандарт пропонує провести їх оцінку, як у кількісному, так і у якісному вимірі. Оцінювання має охоплювати не лише ймовірність виникнення проблеми, але й масштаб можливих наслідків [7]. Особливо підкреслюється необхідність врахування каскадних ефектів – тобто того, як одна проблема може спричинити інші, пов'язані ризики. Це дозволяє побудувати більш повну картину загроз, які може створювати система ШІ у взаємодії з іншими технологіями або соціальними процесами.

Наступним кроком є обробка ризиків, тобто розробка практичних стратегій для їх зменшення. Це може включати перегляд архітектури самої моделі, посилення технічного або організаційного контролю, страхування ризиків або ж свідоме прийняття певного рівня залишкової небезпеки. Головне, щоб вибрані підходи відповідали як характеру загрози, так і загальній стратегії організації щодо етики та безпеки [7].

Завершальним елементом системи управління ризиками є постійний моніторинг і перегляд. Оскільки системи ШІ розвиваються у часі та взаємодіють з динамічним середовищем,

ISO/IEC 23894 наголошує на необхідності регулярної переоцінки ризиків. Це включає встановлення ключових індикаторів ризику (KRI), перевірку ефективності раніше впроваджених заходів і оперативне оновлення стратегій у відповідь на зміни в технологічному або соціальному контексті [7].

Найбільш відкритими до технологій штучного інтелекту є Об'єднані Арабські Емірати (ОАЕ). ОАЕ демонструють швидкий і стратегічний підхід до регулювання штучного інтелекту. Країна прагне не лише використовувати ШІ у своїй економіці, але й стати глобальним лідером у сфері інноваційного управління. У 2017 р. ОАЕ першими у світі призначили міністра штучного інтелекту, підкреслюючи політичну волю до активного розвитку цієї галузі. Відтоді в країні діє Національна стратегія ШІ 2031, що ставить за мету зробити ШІ ключовим інструментом підвищення ефективності урядових послуг, охорони здоров'я, транспорту та освіти [8].

На відміну від багатьох інших юрисдикцій, ОАЕ не приймають детального законодавчого акту, подібного до європейського AI Act. Натомість, держава діє через галузеве регулювання та ініціативи публічно-приватного партнерства. Наприклад, у сфері фінансових послуг Центральний банк ОАЕ та інші регулятори вже впровадили політики використання ШІ, спрямовані на запобігання зловживанням, шахрайству та упередженості в автоматизованих рішеннях [8].

Цікаво, що в ОАЕ також активно діє Dubai International Financial Centre (DIFC) – спеціальна юрисдикція, яка самостійно впроваджує власні цифрові стандарти. У межах ініціативи DIFC AI and Data Protection Guidelines пропонується фреймворк із використанням ШІ, який поєднує етичні принципи, прозорість та відповідальність розробників. ОАЕ, таким чином, рухаються у напрямку «гнучкого регулювання», яке дозволяє адаптуватися до швидких технологічних змін без надмірної бюрократії [8].

Однак відсутність єдиного національного закону щодо ШІ створює ризик фрагментації правового середовища, особливо для міжнародних компаній, які працюють в різних еміратах або секторах. У перспективі ОАЕ можуть розглянути ухвалення більш цілісного законодавства, яке б забезпечило узгодженість підходів у всіх сферах і дало більше впевненості бізнесу щодо вимог до відповідального впровадження технологій штучного інтелекту.

3. Поточна ситуація регулювання штучного інтелекту в Україні

Україна розробила поетапну дорожню карту для впровадження регулювання штучного інтелекту, орієнтуючись на інтеграцію в європейський цифровий простір та імплементацію майбутніх стандартів ЄС, зокрема AI Act. Основною метою документа є створення такого підходу до ШІ, який поєднує захист прав людини, розвиток інноваційної економіки та підвищення міжнародної конкурентоспроможності українських компаній [9].

Документ пропонує bottom-up підхід, тобто рух від м'яких, позазаконодавчих механізмів до поступового впровадження законодавства. На першому етапі (планується два-три роки) передбачається створення регуляторного середовища: тестові механізми, добровільні кодекси поведінки, оцінка ризиків та публікація «Білої книги» – аналітичного документа з рекомендаціями для держави й бізнесу [9]. Це дозволить учасникам ринку адаптуватися до майбутніх вимог без надмірного тиску.

Другий етап передбачає поступову імплементацію положень європейського AI Act, зокрема в момент, коли це стане вимогою для подальшої інтеграції України до ЄС [9]. Планується адаптація найвимогливіших стандартів захисту прав людини, прозорості та етики у використанні ШІ, але з урахуванням специфіки національного ринку. Такий підхід дозволить Україні не лише забезпечити відповідність міжнародним вимогам, але й залишитись гнучкою та конкурентною на глобальному ринку.

У «Білій книзі» окреслено три стратегічні цілі: забезпечення прав людини, підтримка конкурентоспроможності бізнесу та євроінтеграція. Україна прагне гармонізувати майбутнє законодавство із нормами ЄС, аби надати українським ШІ-продуктам вільний доступ до рин-

ку Європи. Водночас у сфері оборони регулювання ШІ не передбачається, з огляду на військовий стан і потребу в інноваціях для захисту держави [10].

Особливу увагу в «Білій книзі» приділено балансу між правами людини та інноваційністю. Міністерство цифрових трансформацій України визнає: надмірне регулювання може загальмувати розвиток ШІ-індустрії, а його повна відсутність – створити загрози для прав і свобод громадян. Тому Україна орієнтується на сервісну модель держави, яка не тисне, а допомагає – через створення публічних інструментів, як-от веб-портал відповідального ШІ, платформа юридичної допомоги, інструмент добровільного маркування систем [10].

Методологія оцінки впливу ШІ на права людини стане базовим інструментом: вона допоможе визначити рівень ризику продукту і стане передумовою для участі в регуляторній пісочниці або отримання консультацій. Регуляторна пісочниця, своєю чергою, дозволить стартапам і компаніям тестувати свої рішення під наглядом держави, готуючись до майбутніх вимог. Також планується система добровільного маркування ШІ, подібна до етикеток на продуктах – аби користувачі знали, як саме працює система, і могли оцінити її надійність, безпеку та етичність.

4. Модель безпеки використання штучного інтелекту для України

Одним із ключових елементів регулювання штучного інтелекту є побудова системи безпеки, яка гарантує відповідальне, етичне та надійне використання ШІ. Для України така модель повинна враховувати як світові стандарти, так і національні виклики, пов'язані з війною, цифровою трансформацією та інтеграцією в європейський правовий простір. У цьому розділі подано узагальнений підхід до формування моделі безпеки використання ШІ в українських умовах, модель наведено на рис. 1.

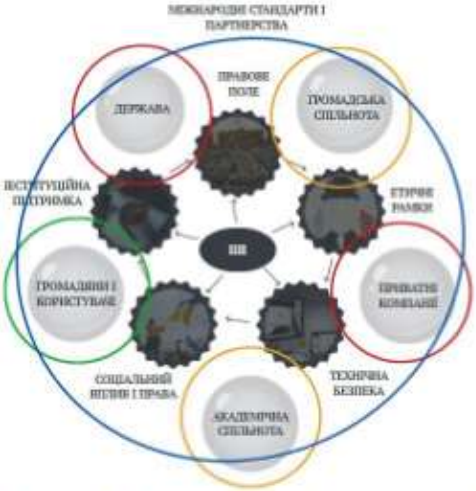

Рис. 1. Модель регулювання безпеки штучного інтелекту в Україні

Запропонована модель безпеки використання штучного інтелекту базується на системному підході, який передбачає взаємодію різних зацікавлених сторін і ключових сфер впливу. У центрі моделі знаходиться ШІ як технологічне ядро, довкола якого формуються міжсекторальні механізми регулювання, етики та технічної підтримки.

У моделі закладено взаємодію між державою, академічною спільнотою, громадянським суспільством, приватними компаніями, міжнародними партнерами та самими користувачами. Такий формат дозволяє не лише регулювати ШІ, а й формувати довіру до нього як до інструменту для розвитку, а не загрози. Україна – країна, що веде активну цифровізацію і не боїться впроваджувати нові технології. У цьому плані український громадський сектор та законодавча база є більш гнучкою і готовою до змін, аніж у сусідніх країнах Європейського Союзу.

Відповідно до рівня ризику, який може виникнути у тій чи іншій складовій її позначено відповідним кольором (де зелений – допустимий, жовтий – високий, червоний – підвищений). За концепцією, схожою з AI Act, цей поділ є необхідним для оцінки кожної моделі ШІ, яка використовується у тій чи іншій галузі. Україна могла б створити «відкритий етичний реєстр ШІ», куди кожна компанія добровільно додала б свою модель. І де кожна модель має власний «паспорт прозорості» – хто її створив, як навчав, які ризики враховані. Додатково кожна система проходила б оцінку на рівень ризику, відповідно до цього її власники мали б забезпечувати певний рівень безпеки, який відповідав би українському стандарту. Це не цензура, це – цифрова культура, адже користувачі мають знати, що саме використовують. Відповідно в уряді має бути створено спеціальний департамент з питань штучного інтелекту, який видавав би ліцензії безпеки та розглядав би і приймав усі потрібні нормативно-правові рішення. Також одним із важливих рішень має бути обов'язкове маркування продуктів та контенту, створеного штучним інтелектом, задля запобігання шахрайства та недоброчесності.

Окрему роль у цій моделі може відігравати громадська спільнота, залучена до моніторингу та аудиту ШІ-систем. Запровадження публічних механізмів контролю, таких як цифрові платформи для фіксації порушень або несправедливих рішень, дозволить забезпечити не лише технічну, а й соціальну безпеку. Водночас державні сервіси – зокрема платформа «Дія» – можуть стати прикладом прозорого, етичного використання ШІ в публічному управлінні, де кожен громадянин має доступ до зрозумілих пояснень рішень, прийнятих алгоритмом.

Ключову роль також відіграє академічна спільнота, яка здатна не лише досліджувати ризики, а й формувати освітню культуру довкола ШІ. Впровадження національного освітнього треку з етики та технологій, інтеграція відповідних курсів у навчальні програми та підтримка молодих дослідників допоможе створити критичну масу спеціалістів, здатних формувати відповідальне майбутнє ШІ в Україні.

В українському контексті приватні компанії не обмежуються впровадженням технологій – вони стають співавторами національної цифрової безпеки. Особливо у сфері штучного інтелекту бізнес має не лише комерційні, а й моральні зобов'язання: перед клієнтами, державою, суспільством і навіть перед власними працівниками. Українські компанії вже довели, що можуть бути лідерами в етичному програмуванні, відкритих кодах, благодійних проєктах та безпечних цифрових рішеннях.

Майбутнє моделі безпеки ШІ в Україні передбачає особливий соціальний договір між державою та бізнесом: не через штрафи чи контроль, а через довіру, сертифікацію, кооперацію та інноваційне партнерство. Компанії, які добровільно дотримуються стандартів прозорості, маркують свої моделі, відкривають алгоритми на ревізію, можуть отримати «цифрову довіру» – умовне етичне маркування, схоже на ISO, але у сфері штучного інтелекту. Водночас важливо підтримати компанії не тільки вимогами, а й ресурсами та середовищем. Наприклад, держава може створити інкубатори відповідального ШІ, де стартапи отримуватимуть доступ до відкритих даних, консультацій з етики, шаблонів політик – без тиску, без формальностей, з орієнтацією на співпрацю.

Ще однією важливою частиною запропонованої моделі безпеки використання ШІ є відкритість до міжнародної співпраці. Це включатиме обмін даними, впровадження міжнародних стандартизацій та регуляторних актів у інтегрованому для України форматі.

У контексті безпеки штучного інтелекту особливу увагу слід приділяти технічним практикам розробки та впровадження ШІ-рішень у приватному секторі, зокрема в компаніях, які створюють системи для використання в соціально чутливих сферах: оборона, фінанси, охорона здоров'я, освіта. Український бізнес сьогодні має справу не лише з комерційними викликами, а й із загрозами кібератак, викрадення моделей, отруєння даних та спробами маніпулювання алгоритмами. Ключовими напрямами технічної безпеки в межах української моделі ШІ можуть стати: безпечне навчання моделей, моніторинг поведінки моделей у

реальному часі, розмежування доступу до компонентів ШІ систем, впровадження explainable AI та контейнери для тестування.

Окремої уваги заслуговує питання інфраструктурної стійкості: більшість українських ІТ-компаній розміщують свої сервіси в хмарі, часто – за кордоном. В умовах воєнних ризиків важливо створити резервні дата-центри, енергонезалежні вузли або використовувати «розподілену відповідальність» за безпеку з партнерами по хмарним сервісам, із обов'язковим аудитом.

Таким чином, запропонована модель безпеки використання ШІ надаватиме простір для розвитку та новаторства, але при цьому увесь процес від створення моделей штучного інтелекту до їх подальшого безпечного використання буде контрольованим та безпечним. Відкритість до міжнародної співпраці тільки підкреслює намір України підтримувати європейські стандарти безпеки та при цьому дасть можливість ділитись власним досвідом з міжнародними партнерами.

Висновки

У результаті проведеного дослідження було проаналізовано міжнародні підходи до регулювання та управління безпекою штучного інтелекту, зокрема стандарти ЄС, США, ISO/IEC та моделі, що впроваджуються в Об'єднаних Арабських Еміратах. Встановлено, що кожна з цих систем має свої сильні сторони, але не є універсальною для застосування в українському контексті. Найбільш прийнятним для України є гібридний підхід, який поєднує сервісну, гнучку модель співрегулювання з орієнтацією на права людини та прозорість, характерну для європейського AI Act.

Особливої уваги потребує побудова власної національної моделі безпеки ШІ, яка враховуватиме специфіку воєнного часу, цифрової трансформації, високої ролі громадянського суспільства та унікального українського досвіду цифрового спротиву. Запропонована модель взаємодії між державою, бізнесом, академічною спільнотою, громадянами та міжнародними партнерами створює основу для формування довіри до ШІ та ефективного управління ризиками на всіх етапах життєвого циклу технологій.

Окрема роль у цій системі належить приватному сектору, який може не лише впроваджувати етичні стандарти, а й бути рушієм цифрової культури безпеки. Важливо, щоб держава підтримувала інноваційні ініціативи через інкубатори відповідального ШІ, сертифікування та механізми прозорості, що спростять адаптацію бізнесу до майбутнього регулювання.

Таким чином, Україна має реальну можливість не лише гармонізувати своє законодавство із міжнародними стандартами, а й запропонувати власний – інноваційний та адаптивний – підхід до безпеки штучного інтелекту. Це дозволить країні стати повноцінним учасником глобального цифрового ринку, водночас зберігаючи пріоритет прав людини, технологічну відкритість та національну стійкість.

Список літератури:

1. Cost of a Data Breach Report 2024. [Електронний ресурс]. Режим доступу: https://www.ibm.com/reports/data-breach.

2. IBM Report: Escalating Data Breach Disruption Pushes Costs to New Highs. [Електронний ресурс]. Режим доступу: https://newsroom.ibm.com/2024-07-30-ibm-report-escalating-data-breach-disruption-pushes-costs-to-new-highs?utm_source=chatgpt.com.

3. Атаки на основі штучного інтелекту: нові виклики для кібербезпеки. [Електронний ресурс]. Режим доступу: https://wezom.com.ua/ua/blog/ataki-na-osnovi-shtuchnogo-intelektu-novi-vikliki-dlya-kiberbezpeki.

4. 5 реальних прикладів хакерських атак за допомогою ШІ. [Електронний ресурс]. Режим доступу: https://dev.ua/news/5-prykladiv-khakerskykh-atak-ai.

5. EU AI Act: first regulation on artificial intelligence. [Електронний ресурс]. Режим доступу: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

6. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. [Електронний ресурс]. Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

7. Information technology – Artificial intelligence – Guidance on risk management [Електронний ресурс]. Режим доступу: https://cdn.standards.iteh.ai/samples/77304/cb803ee4e9624430a5db177459158b24/ISO-IEC-23894-2023.pdf.

8. AI Watch: Global regulatory tracker – United Arab Emirates. [Електронний ресурс]. Режим доступу: https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-uae.

9. Дорожня карта з регулювання штучного інтелекту в Україні Bottom-Up Підхід. [Електронний ресурс]. Режим доступу: https://surli.cc/tyzbug.

10. Біла книга з регулювання ШІ в Україні: бачення Мінцифри. Режим доступу: https://thedigital.gov.ua/storage/uploads/files/page/community/docs/Регулювання%20ШІ.pdf.

Надійшла до редколегії 02.02.2025

Відомості про авторів:

Логачова Єлизавета Олегівна – Харківський національний університет імені В. Н. Каразіна, студентка кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-науковий інститут комп'ютерних наук та штучного інтелекту; Україна; e-mail: lohachova2020kb11@student.karazin.ua; ORCID: https://orcid.org/0000-0002-9815-466X

Єсіна Марина Віталіївна – канд. техн. наук, доцент, Харківський національний університет імені В. Н. Каразіна, в.о. завідувача кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-науковий інститут комп'ютерних наук та штучного інтелекту, АТ Інститут Інформаційних Технологій", науковий співробітник-консультант; Україна; e-mail: m.v.yesina@karazin.ua; ORCID: https://orcid.org/0000-0002-1252-7606

Голубничий Дмитро Юрійович – канд. техн. наук, доцент, АТ "Інститут Інформаційних Технологій", начальник наукового відділу; Україна; ORCID: https://orcid.org/0000-0002-6873-7004

Перші норми, зокрема заборона на використання систем із неприйнятним рівнем ризику, набули чинності вже 2 лютого 2025 р. Інші елементи, як-от кодекси практики та вимоги до прозорості для універсальних моделей ШІ, почнуть діяти через 9 та 12 місяців відповідно [5].

Системи з високим рівнем ризику, включно із тими, що використовуються у сфері охорони здоров'я, освіти, критичної інфраструктури або правоохоронній діяльності, матимуть більше часу на адаптацію. Для них вимоги набудуть чинності через 36 місяців, що дозволить розробникам і користувачам належним чином підготуватися до впровадження нових стандартів.

NIST AI Risk Management Framework – це стратегічний документ, розроблений Національним інститутом стандартів і технологій США (NIST) з метою допомогти організаціям ефективно управляти ризиками, пов'язаними з впровадженням та використанням штучного інтелекту. Він є добровільним, але широко рекомендованим інструментом, який сприяє розробці безпечних, надійних, прозорих і етично обґрунтованих систем ШІ. За цим документом прийнято розподілити можливу нанесену шкоду на три категорії: шкода людям, шкода організації та шкода екосистемі. Шкода людям включає у себе фізичну, психологічну, соціальну або економічну шкоду. Наприклад, фальшиві медичні поради, неправдиві новини, або контент, що провокує насильство. ШІ також може бути використаний для дискримінації, упередженого оцінювання, неправомірного спостереження або цензури – особливо у випадках автоматизованих рішень без належного контролю, що призводить до порушення прав та свободи громадян [6].

Друга категорія під назвою «шкода організації» окреслює такі наслідки, як витіснення творчих професій, або недобросовісну конкуренцію через масове створення фальшивого або дезінформаційного контенту, збої на ринку праці та інших операцій, шкода репутації та інші потенційні ризики безпеці [6].

Третя категорія включає збої у глобальних фінансових системах або системах ланцюга поставок і шкоду навколишньому середовищу та природним ресурсам. Дані категорії створені для запобігання шкодам та ризикам описаним у них, що допомагає сконцентрувати увагу на актуальних проблемах [6].

Документ включає сім основних характеристик для надійних систем штучного інтелекту, яким вони мають відповідати. Першими є валідність та надійність, які означають, що система виконує свою функцію точно, стабільно і в межах запланованого контексту. Вона має бути перевірена на відповідність заявленим цілям і демонструвати передбачувану поведінку навіть в умовах змін середовища чи вхідних даних. Безпечність доповнює цю характеристику, адже система не повинна створювати фізичної або психологічної шкоди користувачам, і має бути захищеною від зловмисного впливу [6].

Неупередженість системи також дуже важлива – це здатність системи уникати дискримінаційних рішень або упередженості. Модель повинна однаково справедливо взаємодіяти з усіма користувачами незалежно від статі, раси, віку чи соціального статусу. Наступною є прозорість – користувачі повинні мати доступ до інформації про те, як система працює, на яких даних вона навчалась, і які можливі обмеження її застосування [6].

Іншою критично важливою характеристикою є пояснюваність – здатність системи та її розробників надати чітке пояснення, чому було прийнято те чи інше рішення. Це особливо актуально у сферах з високим ступенем відповідальності, наприклад, у медицині, фінансах чи юридичній сфері. Також важливою є здатність до захисту конфіденційності, що передбачає дотримання принципів збору, зберігання та обробки персональних даних відповідно до етичних та правових стандартів [6].

Останньою є стійкість – здатність системи зберігати свою функціональність навіть за наявності помилок, атак чи непередбачуваних ситуацій. Надійні системи повинні мати вбудовані механізми відновлення та захисту від зовнішніх загроз [6].

Документ базується на п'яти ключових функціях: Govern, Map, Measure, Manage і Improve [6]. Ці функції утворюють узгоджену систему, яка допомагає організаціям забезпе-

ПРОГРАМА

міжнародної науково-технічної конференції

«КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ В НАУКОЄМНИХ ТЕХНОЛОГІЯХ» (КМНТ-2024)

Харківський національний університет імені В.Н. Каразіна
Координаційна рада НАН України з питань штучного інтелекту
Північного-Східний координаційний науковий центр
з питань штучного інтелекту
Інститут кібернетики імені В.М. Глушкова НАН України
ННЦ Харківський фізико-технічний інститут
Max Planck Institute of microstructure physics
Київський національний університет імені Тараса Шевченка
Національний аерокосмічний університет
імені М. Є. Жуковського «Харківський авіаційний інститут»

Харків-2024

Секція 7

Кібербезпека інформаційних систем і технологій.

Керівник секції: **к.т.н., доцент Єсіна Марина Віталіївна.**

Заст. керівника: **к.т.н., доцент Нарєжній Олексій П.авлович.**

Секретар: **ст. викладач Гальцева Ірина Михайлівна.**

Онлайн засідання секції 7:

28 листопада 2024 р., четвер з 10:00 Часовий пояс: Europe/Kiev

Секція працює за допомогою сервісу Google Meet; посилання для входу:
https://meet.google.com/pjy-cgdm-akh

Номер телефону для приєднання до відеозустрічі:
(US) +1 424-262-6930,
PIN-код: 238 742 074#

1. **АВЕРКОВ О. Ю., ЛИСИЦЬКА І. В.**
   ПРОГРАМНА РЕАЛІЗАЦІЯ ПЕРШОГО ЕТАПУ РОБОТИ ПРОТОКОЛУ ZK-STARK «АРИФМЕТИЗАЦІЯ» ТА РЕЗУЛЬТАТИ ЇЇ ТЕСТУВАНН
2. **БУРЧЕНКО С.Б.**
   АНАЛІЗ ПІДХОДІВ ДО ВИЗНАЧЕННЯ ПРІОРИТЕТІВ КІБЕРІНЦИДЕНТІВ
3. **ГАСІЛІН Д. Л., ЖУРАВЕЛЬ Ю. І.**
   МОДИФІКАЦІЯ МЕТОДУ ПРИХОВУВАННЯ ІНФОРМАЦІЇ НА ОСНОВІ ПЕРЕТВОРЕННЯ КОЛІРНОГО ПРОСТОРУ
4. **ГРІЩЕНКО МИКОЛА РОМАНОВИЧ**
   АНАЛІЗ ВРАЗЛИВОСТЕЙ В БЕЗПЕКІ ВЕБ-ЗАСТОСУНКІВ НА ОСНОВІ АРХІТЕКТУР МІКРОСЕРВІСІВ ТА МЕХАНІЗМИ ЗАХИСТУ
5. **ДОРЕНСЬКИЙ О. П, ІСАЧЕНКОВ Е. В.**
   СТРУКТУРНО-ФУНКЦІОНАЛЬНА МОДЕЛЬ ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ КЛІЄНТСЬКОГО ЗАСТОСУНКУ БАЗИ АНАЛІТИЧНИХ ДАНИХ ҐРУНТІВ
6. **КРАВЧУК Н.В, КОРОБЕЙНІКОВА Т. І.**
   АЛГОРИТМ KNN У БОРОТЬБІ З КІБЕРАТАКАМИ: АНАЛІЗ ВИЯВЛЕННЯ SQL-ІН'ЄКЦІВ
7. **ЛЕВЧЕНКО ДМИТРО ОЛЕКСАНДРОВИЧ**
   АНАЛІЗ ТЕХНОЛОГІЙ ПРОЕКТУВАННЯ ВЕБ ДОДАТКІВ З ІМПЛЕМЕНТАЦІЄЮ МЕТОДІВ ЗАХИСТУ ІНФОРМАЦІЇ
8. **ЛОГАЧОВА Є. О, ЄСІНА М. В..**
   ШТУЧНИЙ ІНТЕЛЕКТ: СУЧАСНІ ЗАГРОЗИ ТА БЕЗПЕКА У
9. **ПАВЛЕНКО Е.О.**
   БАГАТОФАКТОРНА АВТЕНТИФІКАЦІЯ В ІНФОРМАЦІЙНИХ СИСТЕМАХ
10. **СТЕШЕНКО І. Є., ОЛІЙНИКОВ Р. В.**
    АНАЛІЗ І РОЗРОБКА МЕТОДИКИ ВИБОРУ ЗАСОБІВ ПОБУДОВИ СИСТЕМИ МЕРЕЖЕВОЇ БЕЗПЕКИ
11. **ФЕДЧУК Т. Б, КОРОБЕЙНІКОВА Т. І.**
    АЛГОРИТМ РОБОТИ АНАЛІЗАТОРА ПЕРЕНАПРАВЛЕННЯ DOH-ТРАФІКУ
12. **ЯМНИЧ А. Б., КОРОБЕЙНІКОВА Т. І.**
    ДИНАМІЧНІ ПІДХОДИ ДО КЛАСИФІКАЦІЇ ІНФОРМАЦІЇ В СИСТЕМАХ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ПІДПРИЄМСТВ

УДК 004.056.55

Є. О. ЛОГАЧОВА, М. В. ЄСІНА, канд. техн. наук, Д. Ю. ГОЛУБНИЧИЙ, канд. техн. наук

ОЦІНКА ЕФЕКТИВНОСТІ ЗАСТОСУВАННЯ МІЖНАРОДНИХ СТАНДАРТІВ ТА РЕГУЛЯТОРНИХ АКТІВ З РЕГУЛЮВАННЯ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ УКРАЇНИ

Вступ

Наразі технології в Україні стрімко розвиваються. Україна – одна з перших країн у світі, хто просувається на шляху масштабної діджиталізації. Одним із важливих щаблів розвитку у сучасному світі є правильна адаптація до використання технологій штучного інтелекту (ШІ).

Оскільки Україна ще не впровадила активно власні рішення, розглянемо досвід інших країн та спробуємо обрати найкращу концепцію розвитку. Одні країни орієнтовані на впровадження жорстких обмежень, інші роблять усе, щоб не обмежувати розвиток технології.

З метою проведення оцінювання можливостей інтеграції досвіду інших країн у галузі штучного інтелекту для України, було проведено оцінку адекватності кожного з підходів в рамках українського простору. Для оцінювання було використано різні математичні методики, які є універсальними у проведенні багатьох дослідницьких оцінок.

1. Обґрунтування вибору апаратно-програмного забезпечення та методу оцінювання

Для проведення дослідження було обрано удосконалену методику оцінювання для порівняння різних підходів за певними обраними критеріями. Адже дана методика є універсальною і може застосовуватись у різних випадках і для різних методик чи систем.

Необхідними даними для проведення оцінки є обрані критерії, на основі яких буде проведено оцінку експертів, а після – отримано коефіцієнти для порівняння.

Методика ґрунтується на експертному оцінювання за обраними критеріями. Для вибору експертів було враховано їх компетентність у даній галузі. В якості вихідних даних отримуються певні коефіцієнти, що знаходяться за визначеними математичними виразами. Спираючись на дані, отримані від експертів, спочатку здійснюється математична обробка відповідно до правил, визначених у кожному методі оцінювання, а далі – на основі отриманих результатів проводиться програмне моделювання.

В якості середовища для проведення розрахунків та обробки даних було обрано MathCad. Адже саме це середовище містить необхідний пакет інструментів для роботи.

2. Послідовність дій оцінювання

Для проведення оцінювання було виконано наступні дії, які фактично є універсальним алгоритмом для проведення подібних оцінювань [1, 2]:

1) вибір експертів за рівнем їх компетентності;
2) вибір критеріїв (умовних та безумовних) для оцінювання;
3) проведення оцінки за безумовними критеріями;
4) на підставі безумовного інтегрального критерію ухвалюється рішення – залежно від отриманої оцінки – щодо доцільності подальшого порівняння та аналізу;
5) за умови отримання позитивної оцінки – проведення оцінювання за умовними критеріями;
6) опис результатів та висновки.

3. Дослідження за безумовними критеріями та його результати

Для дослідження було обрано декілька різних та часом дещо схожих підходів запропонованих різними країнами або об'єднаннями для регулювання безпеки моделей штучного інтелекту. Для оцінки було відібрано наступні країни та їх підходи: ЄС (AI Act), США (NIST AI RMF), ISO/IEC 23894, Велика Британія (White Paper), ОАЕ, Канада (AIDA), Південна Корея (AI Basic Act), Китай, Албанія. [3-11]

Під безумовними критеріями розуміються ті, що є обов'язковими для виконання. Тобто ті досліджувані підходи, що не відповідають безумовним критеріям не будуть задіяні у оцінці за умовними критеріями.

У якості безумовних критеріїв у нашій оцінці було обрано наступні п'ять показників:

- K_{s1} – валідність;
- K_{s2} – гнучкість;
- K_{s3} – етичність;
- K_{s4} – складність;
- K_{s5} – прогресивність.

Під валідністю мається на увазі реалізованість методу (регуляторного акту чи стандарту) – чи вже впроваджено на практиці, чи тільки заплановано на майбутнє.

Гнучкість – критерій, що описує інтегрованість підходу (регуляторного акту чи стандарту) у різні сфери.

Етичність характеризує те, чи метод справляється із захистом прав людини і чи не порушує підхід усталені норми етики та моралі.

Складність характеризує наявність ускладнень при інтеграції у країні, в нашому контексті – в Україні.

Прогресивність – критерій для визначення того чи сприяє метод нормальному розвитку технології.

Оскільки перелічені часткові критерії належать до безумовних, відбір здійснюється за допомогою бінарної логічної змінної, а саме «так» чи «ні» або 1/0. Таким чином, безумовний критерій може бути представлений у вигляді (1) [1, 2]:

$$(K_{s1}, K_{s2}, K_{s3}, K_{s4}, K_{s5}) \in (1,0). \quad (1)$$

Враховуючи перераховані безумовні критерії та (1), функція відповідності адекватності методу (регуляторного акту чи стандарту) може бути записана у вигляді (2) [1, 2]:

$$f_{\text{фв}}() = (K_{s1} \wedge K_{s2} \wedge K_{s3} \wedge K_{s4} \wedge K_{s5}) = K_s. \quad (2)$$

Отже, якщо $f_{\text{фв}} = 1$ значить метод відповідає вимогам. Ті методи, що відповідають вимогам будуть розглянуті далі [1, 2]. Результати порівняння відносно безумовних критеріїв наведено у табл. 1.

Таблиця 1 – Результати порівняння відносно безумовних критеріїв

Критерій / Метод	K_{s1}	K_{s2}	K_{s3}	K_{s4}	K_{s5}	K_s
ЄС (AI Act)	1	1	1	1	1	1
США (NIST AI RMF)	1	1	1	1	1	1
ISO/IEC 23894	1	1	1	1	1	1
Велика Британія (White Paper)	1	1	1	1	1	1
ОАЕ	0	1	0	0	1	0
Канада (AIDA)	1	1	1	1	1	1
Південна Корея (AI Basic Act)	1	1	1	1	1	1
Китай	1	1	1	1	1	1
Албанія	0	1	0	0	1	0

За отриманими результатами, негативну оцінку отримали підходи, що описані у документах Албанії та ОАЕ [6,7]. Усі інші будуть досліджені далі на основі умовних критеріїв та інтегрального умовного критерію ,та потенційно розглянуті, як можливі підходи, що доступні для застосування в Україні.

4. Оцінювання за допомогою методу попарних порівнянь

В якості умовних критеріїв було обрано часткові критерії наведені у табл. 2. Наступним кроком було побудовано табл. 3, в якій наведено оцінку вкладу кожного критерію.

Під час парного порівняння експерт оцінює об'єкти попарно, визначаючи, який із двох є важливішим. Усі можливі комбінації об'єктів подаються у вигляді матриці попарних порівнянь. Під час проведення парних порівнянь необхідно відповісти на запитання: який із двох елементів є важливішим, має більший вплив, вищу ймовірність або перевагу. Зазвичай при порівнянні критеріїв визначають, який із них є значущішим відносно одне одного [1, 12].

Таблиця 2 – Умовні часткові критерії оцінки ЕП

№	Критерії	Позначення
1	Рівень прозорості алгоритмів і систем прийняття рішень, тобто пояснюваність процедур, вимог та інших дій зазначених у документі.	K_{y1}
2	Перспективність інтеграції регуляторного акту, методу, стандарту в Україні.	K_{y2}
3	Рівень гнучкості документу у застосуванні в різних сферах.	K_{y3}
4	Можливість запобігання виникненню етичних ризиків, пов'язаних з порушенням прав та свобод людини.	K_{y4}
5	Рівень інституційної підтримки та державної залученості.	K_{y5}
6	Вплив на перспективність розвитку технології окремо та у інших технологічних галузях.	K_{y6}

Для оцінки кожного критерію було побудовано матрицю попарних порівнянь.

Таблиця 3 – Попарне порівняння критеріїв для визначення важливості

Критерії	K_{y1}	K_{y2}	K_{y3}	K_{y4}	K_{y5}	K_{y6}	q_i	r_i
K_{y1}	1	1	2	0,5	4	2	1,41421356	0,18940253
K_{y2}	1	1	1	0,33	2	3	1,12246205	0,15032889
K_{y3}	0,5	1	1	0,2	3	1	0,81818882	0,10957824
K_{y4}	2	3	6	1	6	3	2,94168275	0,39397315
K_{y5}	0,25	0,5	0,33	0,1667	1	0,1428	0,31580809	0,04229549
K_{y6}	0,5	0,33	1	0,33	7	1	0,85435347	0,1144217

Відношення узгодженості має значення 7,466708745.

Таким чином було отримано розуміння того, який з критеріїв буде відігравати найбільшу роль при оцінюванні.

Далі було проведено окреме оцінювання за певним критерієм, приклад такого оцінювання наведено у табл. 4.

Таблиця 4 – Матриця попарних порівнянь по критерію K_{y1}

Методи (документи)	ЄС (AI Act)	США (NIST AI RMF)	ISO/IEC 23894	Велика Британія (White Paper)	Канада (AIDA)	Південна Корея (AI Basic Act)	Китай	q_i	r_i
ЄС (AI Act)	1	1	1	2	2	2	3	1,574610106	0,20586303
США (NIST AI RMF)	1	1	1	2	2	2	3	1,574610106	0,20586303
ISO/IEC 23894	1	1	1	2	2	2	3	1,574610106	0,20586303
Велика Британія (White Paper)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Канада (AIDA)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Південна Корея (AI Basic Act)	0,5	0,5	0,5	1	1	1	2	0,820335356	0,10724986
Китай	0,333333	0,3333333	0,3333333	0,5	0,5	0,5	1	0,463987849	0,06066133

Розрахунки вектору пріоритетів було проведено на основі добутку вектору пріоритетів першого рівня, отриманого раніше у табл. 4, та матриці отриманих значень.

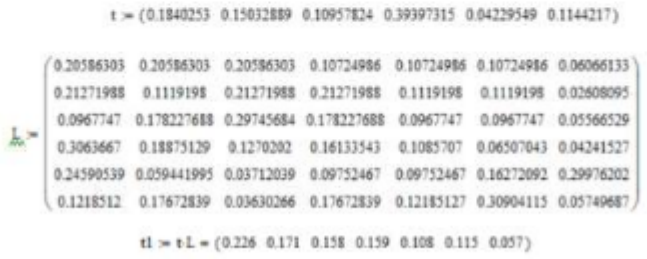

Рисунок 1 – Розрахунок результуючого вектору пріоритетів

В результаті отримано чисельні дані, за якими найкращий результат показав європейський підхід – ЄС (AI Act) – 0,226, а найгірші значення у контексті обраних критеріїв отримав китайський підхід до регулювання ШІ – 0,057. Значення інших розглянутих підходів налічують: США (NIST AI RMF) – 0,171; ISO/IEC 23894 – 0,158; Велика Британія (White Paper) – 0,159; Канада (AIDA) – 0,108; Південна Корея (AI Basic Act) – 0,115.

5. Оцінювання методом визначення вагових коефіцієнтів за допомогою шкали Фішберна

Даний метод полягає у тому, що за основу беруться оцінки певної кількості експертів, у нашому випадку їх було п'ять. Експерти оцінювали важливість того чи іншого критерію ранжуючи їх – табл. 5.

Таблиця 5 – Ранжування критеріїв експертами

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	2	1	4	3	5	6
2	5	1	3	4	2	6
3	3	2	4	2	6	5
4	5	1	4	2	6	3
5	3	2	5	1	6	4

Після було проведено розрахунки для визначення значень вагових коефіцієнтів за формулою [1, 2, 13]:

$$a_i = \frac{2\cdot(m-i+1)}{m\cdot(m+1)},$$

де a_i – ваги i-го показника, i – номер показника, m – кількість показників.

Результати визначення значень вагових коефіцієнтів для критеріїв наведено у табл. 6.

Таблиця 6 – Значення вагових коефіцієнтів та їх середнє значення

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,23809524	0,28571429	0,14285714	0,19047619	0,0952381	0,04761905
2	0,0952381	0,28571429	0,19047619	0,14285714	0,23809524	0,04761905
3	0,19047619	0,23809524	0,14285714	0,28571429	0,04761905	0,0952381
4	0,0952381	0,28571429	0,14285714	0,23809524	0,04761905	0,19047619
5	0,19047619	0,23809524	0,0952381	0,28571429	0,04761905	0,14285714
w_i	0,16190476	0,26666667	0,14285714	0,22857143	0,0952381	0,1047619

Далі також було проведено такі ж дії для кожного з підходів окремо. Приклад розрахунків наведено у табл. 7.

Таблиця 7 – Значення вагових коефіцієнтів та їх середнє значення для ЄС (AI Act)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	0,238095238	0,28571429	0,19047619	0,14285714	0,04761905	0,0952381
2	0,285714286	0,19047619	0,14285714	0,23809524	0,04761905	0,0952381
3	0,19047619	0,23809524	0,0952381	0,28571429	0,04761905	0,14285714
4	0,285714286	0,23809524	0,14285714	0,19047619	0,04761905	0,0952381
5	0,238095238	0,28571429	0,0952381	0,19047619	0,14285714	0,04761905
w_i	0,247619048	0,24761905	0,13333333	0,20952381	0,06666667	0,0952381

Надалі було проведено саме оцінювання та отримано результати – рис. 2. Найкращий результат отримав канадський підхід Канада (AIDA) – 0,267. А найгірший результат отримав США (NIST AI RMF) – 0,162.

Інші підходи отримали наступні показники: ЄС (AI Act) – 0,247; китайський підхід – 0,168; ISO/IEC 23894 – 0,185; Велика Британія (White Paper) – 0,185; Південна Корея (AI Basic Act) – 0,191.

$w1 := (0.1619048 \;\; 0.26666667 \;\; 0.14285714 \;\; 0.22857143 \;\; 0.095231 \;\; 0.1047619)$

$$L\_Fish := \begin{pmatrix} 0.247619048 & 0.24761905 & 0.13333333 & 0.20952318 & 0.6666667 & 0.0952381 \\ 0.142857143 & 0.1714257 & 0.1428514 & 0.1428514 & 0.2 & 0.2 \\ 0.26666667 & 0.1904619 & 0.11428571 & 0.24761905 & 0.05714286 & 0.12380952 \\ 0.2 & 0.18095238 & 0.21904762 & 0.25714286 & 0.0952381 & 0.0461905 \\ 0.21904619 & 0.2 & 0.20952381 & 0.22857143 & 0.952381 & 0.0461905 \\ 0.20952318 & 0.20952318 & 0.15238095 & 0.28571429 & 0.0461905 & 0.0952381 \\ 0.24619048 & 0.12380952 & 0.07619048 & 0.24761905 & 0.21904762 & 0.06666667 \end{pmatrix}$$

$V\_Fish := w1$

$V\_Fish = (0.162 \;\; 0.267 \;\; 0.143 \;\; 0.229 \;\; 0.095 \;\; 0.105)$

$Fish\_rez := L\_Fish \cdot V\_Fish^{T}$

$Fish\_rez^{T} = (0.247 \;\; 0.162 \;\; 0.185 \;\; 0.185 \;\; 0.267 \;\; 0.191 \;\; 0.168)$

Рисунок 2 – Обчислення результуючого вектору пріоритетів

6. Оцінювання методом визначення вагових коефіцієнтів на основі методу ранжування

Припустимо, що існує m часткових показників і n експертів, які оцінюють їхню важливість для певної системи. У нашому випадку це знову ж п'ять експертів. Найвищий ранг (оцінку) m отримує показник, який вважається найважливішим, далі – показники з меншими рангами відповідно до зниження значущості, а ранг 1 присвоюється найменш важливому показнику. У цьому разі вагові коефіцієнти розраховуються за формулою [1, 2, 14]:

$$w_j = \frac{r_j}{\sum_{j=1}^{m} r_j}.$$

Як і у попередніх випадках, спершу було проведено розрахунки вагових коефіцієнтів для критеріїв – табл. 8.

Таблиця 8 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	6	5	2	4	3	1
2	5	4	1	6	3	2
3	6	4	3	5	2	1
4	5	3	2	6	4	1
5	6	4	2	5	1	3
$r_j = \sum_{i=1}^{n} r_{ij}$	28	20	10	26	13	8
w_j	0,26666667	0,19048	0,09524	0,24762	0,12381	0,07619

Також у наступних кроках було розраховано вагові коефіцієнти для кожного з підходів, які ми розглянули. Приклад такого розрахунку наведено у табл. 9.

Далі за схожим принципом розрахунку результуючого вектору було проведено розрахунки та отримано результат – рис. 3. За яким: найкращий показник отримав європейський підхід – документ AI Act – 0,197. Найгірші результати показали Велика Британія White Paper – 0,176 та США NIST AI RMF – 0,177. Інші підходи отримали оцінки: ISO/IEC 23894 – 0,186; китайський підхід – 0,183; Південна Корея (AI Basic Act) – 0,187; Канада (AIDA) – 0,195.

Таблиця 9 – Значення вагових коефіцієнтів для Велика Британія (White Paper)

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6
1	3	5	6	4	2	1
2	4	6	5	3	1	2
3	3	6	5	4	2	1
4	4	6	5	3	2	1
5	3	6	5	4	1	2
$r_j = \sum_{i=1}^{n} r_{ij}$	17	29	26	18	8	7
w_j	0,16190476	0,27619	0,24762	0,17143	0,07619	0,06667

$$w2 := (0.26666667\ \ 0.19048\ \ 0.09524\ \ 0.24762\ \ 0.12381\ \ 0.07619)$$

$$L\_Koef := \begin{pmatrix} 0.21904762 & 0.19048 & 0.12381 & 0.28571 & 0.12381 & 0.05714 \\ 0.16190476 & 0.2851 & 0.2381 & 0.17143 & 0.06667 & 0.07619 \\ 0.28571429 & 0.20952 & 0.21905 & 0.13333 & 0.09524 & 0.05714 \\ 0.16190476 & 0.27619 & 0.24762 & 0.17143 & 0.07619 & 0.06667 \\ 0.20952381 & 0.21905 & 0.14286 & 0.2851 & 0.0619 & 0.06667 \\ 0.17142857 & 0.1619 & 0.0619 & 0.27619 & 0.24762 & 0.06667 \\ 0.18095238 & 0.14286 & 0.08571 & 0.2381 & 0.27619 & 0.07619 \end{pmatrix}$$

$$V\_Koef := w2$$

$$V\_Koef = (0.267\ \ 0.19\ \ 0.095\ \ 0.248\ \ 0.124\ \ 0.076)$$

$$Koef\_rez := L\_Koef \cdot V\_Koef^{T}$$

$$Koef\_rez^{T} = (0.197\ \ 0.177\ \ 0.186\ \ 0.176\ \ 0.195\ \ 0.187\ \ 0.183)$$

Рисунок 3 – Обчислення результуючого вектору пріоритетів

7. Оцінювання методом визначення вагових коефіцієнтів на основі методу приписування балів

У методі приписування балів маємо m показників, важливість яких оцінює n експертів.

Кожен експерт визначає, наскільки важливим є кожен показник для досліджуваної системи, і виставляє йому оцінку за шкалою від 0 до 10.

Допускається використання дробових значень, якщо експерт вважає, що показник заслуговує проміжної оцінки. Також кілька показників можуть отримати однаковий бал, якщо їх важливість вважається рівною.

Наступним кроком розраховується вага кожного коефіцієнта, використовуючи формули [1, 2]:

$$r_{ij} = \frac{h_{ij}}{\sum_{j=1}^{m} h_{ij}},\ \sum_{j=1}^{m} r_{ij} = 1,$$

де r_{ij} – ваги j-го показника, визначені i-м експертом; h_{ij} – бал i-го експерта, виставлений j-му показнику, n – кількість експертів, m – кількість показників.

В результаті вагові коефіцієнти визначались за формулами [1, 2]:

$$w_j = \frac{\sum_{i=1}^{n} r_{ij}}{\sum_{j=1}^{m}\sum_{i=1}^{n} r_{ij}};\ \sum_{j=1}^{n} w_j = 1.$$

Розрахунок вагових коефіцієнтів для критеріїв наведено у табл. 10.

Таблиця 10 – Значення вагових коефіцієнтів

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	8	5	10	7	6	45	0,2	0,17778	0,11111	0,22222	0,15556	0,13333
2	7	10	5	9	8	6	45	0,15555556	0,22222	0,11111	0,2	0,17778	0,13333
3	6	7	9	8	5	10	45	0,13333333	0,15556	0,2	0,17778	0,11111	0,22222
4	8	6	4	10	7	5	40	0,2	0,15	0,1	0,25	0,175	0,125
5	6	9	5	10	8	7	45	0,13333333	0,2	0,11111	0,22222	0,17778	0,15556
							$\sum_{i=1} r_j$	0,82222222	0,90555556	0,63333333	1,07222222	0,79722222	0,76944444
							w_j	0,16444444	0,18111111	0,12666667	0,21444444	0,15944444	0,15388889

Далі було проведено такі ж розрахунки для кожного з підходів, приклад наведено у табл. 11.

Таблиця 11 – Значення вагових коефіцієнтів для ISO/IEC 23894

Показники / Експерти	x_1	x_2	x_3	x_4	x_5	x_6	$\sum_{j=1}^{m} h_{ij}$	Ваги показників					
								r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}
1	9	8	9	7	7	8	48	0,1875	0,16666667	0,1875	0,14583333	0,14583333	0,16666667
2	10	8	9	6	7	7	47	0,21276596	0,17021277	0,19148936	0,12765957	0,14893617	0,14893617
3	8	7	8	7	6	7	43	0,18604651	0,1627907	0,18604651	0,1627907	0,13953488	0,1627907
4	9	7	8	7	6	7	44	0,20454545	0,15909091	0,18181818	0,15909091	0,13636364	0,15909091
5	9	8	8	6	6	8	45	0,2	0,17777778	0,17777778	0,13333333	0,13333333	0,17777778
							$\sum_{i=1} r_j$	0,99085792	0,83653882	0,92463183	0,72870785	0,70400136	0,81526222
							w_j	0,19817158	0,16730776	0,18492637	0,14574157	0,14080027	0,16305244

В результаті було обчислено значення результуючого вектору пріоритетів – рис. 4. Та отримано результат оцінювання: найкращий результат за розрахунками отримали ЄС (AI Act) та Велика Британія (White Paper) – 0,169. Найгірший результат показав китайський метод – 0,161. Хоча варто помітити, що розходження між усіма методами – мінімальні: США (NIST AI RMF) – 0,167; ISO/IEC 23894 – 0,164; Канада (AIDA) – 0,163; Південна Корея (AI Basic Act) – 0,162.

$$w3 := (0.16444444 \;\; 0.18111111 \;\; 0.12666667 \;\; 0.21444444 \;\; 0.1544444 \;\; 0.15388889)$$

$$I\_Point = \begin{pmatrix} 0.18559759 & 0.19759419 & 0.12088985 & 0.1895959 & 0.16525484 & 0.141006593 \\ 0.18368019 & 0.18309992 & 0.16125155 & 0.17913884 & 0.13583563 & 0.15719388 \\ 0.19817158 & 0.16730776 & 0.18492657 & 0.14574157 & 0.14080027 & 0.16305244 \\ 0.20007323 & 0.1744411 & 0.14047564 & 0.19998073 & 0.13187249 & 0.1531568 \\ 0.16796373 & 0.17222652 & 0.20166662 & 0.13462316 & 0.1301903 & 0.19332967 \\ 0.16244218 & 0.15809436 & 0.19657724 & 0.12821462 & 0.20508788 & 0.14958372 \\ 0.13253901 & 0.15516997 & 0.18277929 & 0.10947515 & 0.21906104 & 0.20097554 \end{pmatrix}$$

$$V\_Point := w3$$

$$V\_Point = (0.164 \;\; 0.181 \;\; 0.127 \;\; 0.214 \;\; 0.154 \;\; 0.154)$$

$$Point\_rez := I\_Point \, V\_Point^{T}$$

$$Point\_rez^{T} = (0.169 \;\; 0.167 \;\; 0.164 \;\; 0.169 \;\; 0.163 \;\; 0.162 \;\; 0.161)$$

Рисунок 4 – Обчислення результуючого вектору пріоритетів

8. Оцінювання методом визначення вагових коефіцієнтів на основі числового способу

Останнім застосованим методом став метод визначення вагових коефіцієнтів на основі числового способу. У цьому методі для кожного показника обчислюється коефіцієнт відносного розкиду за формулою [1, 2]:

$$\delta_i = \frac{x_{i\max} - x_{i\min}}{x_{i\max}}.$$

Оскільки значення самих показників можна було отримати з будь-якого попереднього обчислення, були обрані значення обчислень за шкалою Фішберна. А після проведення обчислення вагових коефіцієнтів за формулою [1, 2]:

$$w_i = \frac{\delta_i}{\sum_{i=1}^{m} \delta_i}.$$

Розрахунки вагових коефіцієнтів для критеріїв наведено у табл. 12.

Таблиця 12 – Значення вагових коефіцієнтів

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,0952381	0,23809524	0,0952381	0,1428571	0,0476191	0,0476191
$x_{i\max}$	0,23809524	0,28571429	0,1904762	0,2857143	0,2380952	0,1904762
δ_i	0,59999998	0,16666667	0,5	0,5	0,8	0,75
w_i	0,18090452	0,05025126	0,1507538	0,1507538	0,241206	0,2261307

Далі, як і у попередніх методах, було повторено розрахунки для усіх досліджуваних підходів регулювання ШІ. Приклад розрахунків наведено у табл. 13. А після проведено обчислення результуючого вектору пріоритетів – рис. 5. У результаті, за даним методом отримано результати: найкращий результат за розрахунками отримали китайський підхід – 0,385 та ЄС (AI Act) – 0,18. Найгірший результат показав Велика Британія (White Paper) – 0,048. Інші методи отримали наступні результати: США (NIST AI RMF) – 0,156; ISO/IEC 23894 – 0,156; Канада (AIDA) – 0,059; Південна Корея (AI Basic Act) – 0,111.

Таблиця 13 – Значення вагових коефіцієнтів для Канади (AIDA)

Показники / Оцінювання	x_1	x_2	x_3	x_4	x_5	x_6
$x_{i\min}$	0,19047619	0,14285714	0,1428571	0,1428571	0,0952381	0,0476191
$x_{i\max}$	0,23809524	0,23809524	0,2857143	0,2857143	0,0952381	0,0476191
δ_i	0,2	0,40000002	0,5	0,5	0	0
w_i	0,125	0,08450705	0,1056338	0,1056338	0	0

$$w4 := (0.18090452 \quad 0.05025126 \quad 0.1507538 \quad 0.1507528 \quad 0.241206 \quad 0.2261307)$$

$$L\_Num := \begin{pmatrix} 0.11111111 & 0.11111112 & 0.1666667 & 0.1666667 & 0.2222222 & 0.2222222 \\ 0.16901409 & 0.17605634 & 0.1690141 & 0.1690141 & 0.1760563 & 0.1408451 \\ 0.06944445 & 0.16666668 & 0.2777778 & 0.1388889 & 0.2083333 & 0.1388889 \\ 0.08450704 & 0.08450705 & 0.084507 & 0.1056338 & 0 & 0 \\ 0.125 & 0.08450705 & 0.1056338 & 0.1056338 & 0 & 0 \\ 0.23529411 & 0.47058824 & 0.2941177 & 0 & 0 & 0 \\ 0.15151515 & 0.39215679 & 0.5882353 & 0.3921569 & 0.2352941 & 0.5882353 \end{pmatrix}$$

$$V\_Num := w4$$

$$V\_Num = (0.181 \quad 0.05 \quad 0.151 \quad 0.151 \quad 0.241 \quad 0.226)$$

$$Num\_rez := L\_Num \cdot V\_Num^{T}$$

$$Num\_rez^{T} = (0.18 \quad 0.165 \quad 0.165 \quad 0.048 \quad 0.059 \quad 0.111 \quad 0.385)$$

Рисунок 5 – Обчислення результуючого вектору пріоритетів

9. Висновки за результатами оцінювання

Провівши оцінювання адекватності різних міжнародних підходів регулювання штучного інтелекту щодо застосування в Україні, було отримано результати, які схематично наведені на рис. 6.

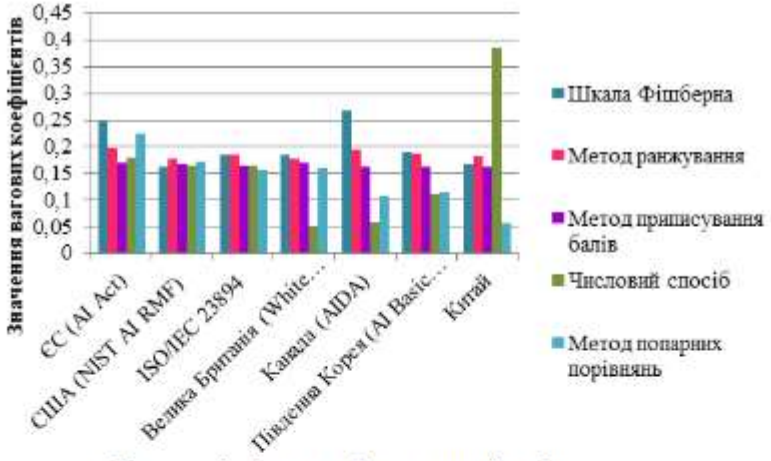

Рисунок 6 – Ілюстрація результатів оцінювання

За результатами описаними раніше та наведеними на рис. 6, видно, що найкраще за різними методами оцінювання виявились ЄС (AI Act) та Канада (AIDA). При чому ЄС (AI Act) показує доволі стабільні значення, що свідчить про збалансованість та зрілість регуляторної системи, де враховуються як прозорість алгоритмів, так і етична складова та інституційна підтримка.

Китайський підхід до регулювання штучного інтелекту показав найгірші результати майже за всіма методиками, тобто він є одним із найменш ймовірно інтегрованих для України. Хоча за аналізом усереднених значень, наведеним у табл. 9.1 та рис. 7, цей підхід показав другий результат.

США (NIST AI RMF) та ISO/IEC 23894 мають близькі результати, з незначною перевагою першого. Обидва документи роблять акцент на гнучкості й стандартизації. Це робить їх доволі хорошими прикладами для України.

Велика Британія демонструє середні показники за всіма методами, що відповідає її гнучкому, ринково-орієнтованому підходу, заснованому на принципах «регулювання через довіру». Цей підхід забезпечує помірну ефективність, але нижчу централізованість.

Південна Корея має дещо нижчі значення, однак показує відносну збалансованість, але все ж це робить її не найкращим прикладом для наслідування у випадку України.

Усереднене оцінювання результатів для кожного з підходів наведено у табл. 14.

Таблиця 14 – Усереднене оцінювання результатів

Методи (документи)	Середній результат
ЄС (AI Act)	0,2038
США (NIST AI RMF)	0,1684
ISO/IEC 23894	0,1716
Велика Британія (White Paper)	0,1474
Канада (AIDA)	0,1584
Південна Корея (AI Basic Act)	0,1532
Китай	0,1908

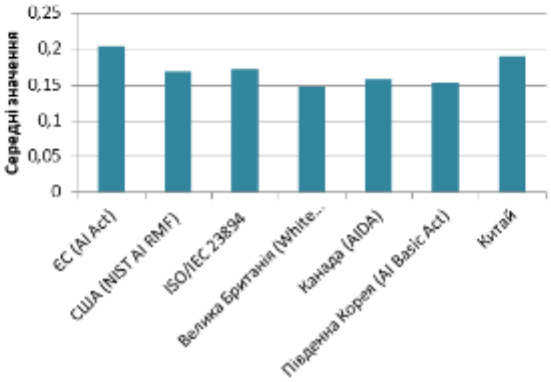

Рисунок 7 – Усереднене оцінювання результатів

Висновки

У результаті дослідження і проведення оцінювання ми дійшли висновків, що одними з найкращих прикладів регулювання ШІ для України є ЄС (AI Act) та Канада (AIDA). Адже саме вони показали найстабільніші позитивні результати.

А серед прикладів, які навряд знайдуть успішний шлях інтеграції в Україні є китайський підхід, який є хоч і доволі цікаво побудованим та все ж акцентує увагу більше не на пріоритетних для України критеріях.

Список літератури:

1. Горбенко І. Д. Методи, методика та результати порівняльного аналізу електронних підписів згідно ДСТУ ISO/IEC 14888-3:2014 / І. Д. Горбенко, М. В. Єсіна // Вісник Національного університету "Львівська політехніка": серія "Автоматика, вимірювання та керування". – Л. : Національний університет "Львівська політехніка", 2016. – № 852 – С. 9–22.

2. Горбенко Ю. І., Ганзя Р. С., Акользіна О. С. Електронні підписи на основі ідентифікаторів та бінарного відображення // Прикладна радіоелектроніка. Х. : Харківський національний університет радіоелектроніки, 2015. Т. 14. № 4. С. 284–290.

3. EU AI Act: first regulation on artificial intelligence. [Електронний ресурс]. – Режим доступу: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

4. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. [Електронний ресурс]. – Режим доступу: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

5. Information technology – Artificial intelligence – Guidance on risk management [Електронний ресурс]. – Режим доступу: https://cdn.standards.iteh.ai/samples/77304/cb803ee4e9624430a5db177459158b24/ISO-IEC-23894-2023.pdf.

6. AI Watch: Global regulatory tracker – United Arab Emirates. [Електронний ресурс]. – Режим доступу: https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-uae.

7. Albania turns to AI to beat corruption and join EU [Електронний ресурс]. – Режим доступу: https://www.politico.eu/article/albania-use-ai-artificial-intelligenve-join-eu-corruption/

8. Interim Measures for the Management of Generative Artificial Intelligence Services [Електронний ресурс]. – Режим доступу: https://www.chinalawtranslate.com/en/generative-ai-interim/#gsc.tab=0

9. A pro-innovation approach to AI regulation 3 August 2023 [Електронний ресурс]. – Режим доступу: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper?utm_source=chatgpt.com#part-7-conclusion-and-next-steps

10. The Artificial Intelligence and Data Act (AIDA) – Companion document [Електронний ресурс]. – Режим доступу: https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document

11. South Korea Artificial Intelligence (AI) Basic Act [Електронний ресурс]. – Режим доступу: https://www.trade.gov/market-intelligence/south-korea-artificial-intelligence-ai-basic-act

12. Oletsky O. Enhancing Consistency of Pairwise Comparisons on the Base of Linear Algebraic Equations. NaUKMA Research Papers. Computer Science. 2023. T. 5. C. 85–91. [Електронний ресурс]. – Режим доступу: http://nrpcomp.ukma.edu.ua/article/view/275238

13. Lyashenko, E. M., Klymash, M. P., Lyashko, O. A. On creation of expert knowledge base for estimation of state of organizational systems // International Journal of Fuzzy Systems and Advanced Applications. – 2014. – Vol. 1. – P. 7–13.

14. Alguliyev, R. M., Nabibayova, G. Ch., & Abdullayeva, S. R. Evaluation of websites by many criteria using the algorithm for pairwise comparison of alternatives // I.J. Intelligent Systems and Applications. – 2020. – Vol. 12, No. 6. – P. 64–74. https://doi.org/10.5815/ijisa.2020.06.05.

ISSN 2519-2310 (Online)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В.Н.КАРАЗІНА
V.N. KARAZIN KHARKIV NATIONAL UNIVERSITY

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА
COMPUTER SCIENCE AND CYBERSECURITY
(CS&CS)

№ 1(27) 2025
Issue 1(27) 2025

Заснований 2015 року

Міжнародний електронний науково-теоретичний журнал
International electronic scientific journal

ISSN 2519-2310 (Online)

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА, 2025, № 1 (27)

COMPUTER SCIENCE AND CYBERSECURITY, 2025, Iss. 1 (27)

DOI: https://doi.org/10.26565/2519-2310-2025-1-05
УДК 340.137

ДОСЛІДЖЕННЯ ТА ПОРІВНЯННЯ НОРМАТИВНИХ РЕГУЛЯТОРНИХ ДОКУМЕНТІВ ЄВРОПЕЙСЬКОГО СОЮЗУ У СФЕРІ КІБЕРБЕЗПЕКИ

Марина Єсіна[1,2], канд. техн. наук, доцент, в.о. завідувача кафедри кібербезпеки інформаційних систем, мереж і технологій, ННІ КШ та ШІ; науковий співробітник-консультант, e-mail: m.v.yesina@karazin.ua, ORCID: https://orcid.org/0000-0002-1137-2382

Єлизавета Логачова[1], студентка 5 курсу кафедри кібербезпеки інформаційних систем, мереж і технологій, ННІ КШ та ШІ, e-mail: lohachova2020kb11@student.karazin.ua, ORCID: https://orcid.org/0000-0002-9815-466X

Євгенія Колованова[1], канд. техн. наук, в.о. заступника директора з навчальної роботи, ННІ КШ та ШІ, e-mail: e.kolovanova@karazin.ua, ORCID: https://orcid.org/0000-0002-0326-2394

[1]Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків, 61022, Україна

[2]ПрАТ "Інститут Інформаційних Технологій", вул. Отамановського професора, 15, Харків, 61022, Україна

Рукопис надійшов 1 березня 2025 р. Отримано після рецензування 1 квітня 2025 р. Прийнято 2 травня 2025 р.

Анотація: У статті представлено ґрунтовне дослідження та порівняння основних нормативно-правових актів Європейського Союзу у сфері кібербезпеки, серед яких: Директива NIS2, Загальний регламент про захист даних (GDPR), Регламент про цифрову операційну стійкість (DORA) та стандарт PCI DSS. Ці документи розглядаються як основоположні елементи формування сучасної політики ЄС у сфері захисту цифрової інформації, що охоплює персональні дані, фінансову інформацію та критичну інфраструктуру. У роботі окреслено ключові завдання кожного з актів, проаналізовано їхню сферу застосування, вимоги до суб'єктів регулювання, механізми управління ризиками, звітування про інциденти, взаємодію з постачальниками, а також санкційні положення. Особливу увагу приділено порівнянню актів за впливом на бізнес та IT-інфраструктуру, а також виявленню взаємозв'язків між ними. Встановлено, що хоча кожен документ має власний фокус - захист персональних даних, стійкість фінансової інфраструктури, безпека цифрових мереж чи захист транзакцій з платіжними картками – всі вони спрямовані на створення цілісної екосистеми кіберзахисту в межах ЄС. Стаття також аналізує міжнародні аналоги згаданих актів, такі як GDPR-подібні закони у США та Бразилії, стандарти NIST та ISO, що свідчить про глобальний характер проблеми цифрової безпеки та пошук спільних підходів до її вирішення. У підсумку, робота підкреслює важливість комплексного та гармонізованого підходу до кібербезпеки як ключової умови сталого розвитку цифрового суспільства. З огляду на актуальні загрози, зокрема геополітичні конфлікти та зростання масштабів кіберзлочинності, ефективна імплементація європейських стандартів набуває особливого значення для країн-партнерів, зокрема України, яка потребує адаптації відповідних норм для підвищення національної кіберстійкості.

Ключові слова: *нормативні документи, європейський союз, регуляторні документи, кібербезпека, бізнес, захист інформації*

ISSN 2519-2310 (Online) Computer Science and Cybersecurity. 2025. Issue 1 (27)

Як цитувати: Єсіна М., Логачова Є., Колованова Є. Дослідження та порівняння нормативних регуляторних документів європейського союзу у сфері кібербезпеки. *Комп'ютерні науки та кібербезпека.* 2025; № 1(27): С. 60–72. https://doi.org/10.26565/2519-2310- 2025-1-05

In cites: Yesina M., Lohachova Ye., Kolovanova Ye. (2025). Research and comparison of European Union regulatory documents in cybersecurity. *Computer Science and Cybersecurity. 1*(27): 60–72. https://doi.org/10.26565/2519-2310- 2025-1-05 (in Ukrainian)

1. Вступ

У сучасному світі цифрові технології стали невід'ємною складовою суспільного, економічного та політичного життя. Разом із безпрецедентними можливостями, які вони відкривають для розвитку економік та суспільств, вони також створюють нові загрози – від кібершахрайства до масових атак на критичну інфраструктуру. У цьому контексті питання кібербезпеки набуло особливої ваги як на національному, так і на міжнародному рівнях. Європейський Союз, як одне із провідних інтеграційних об'єднань у світі, визначає кібербезпеку як пріоритетний напрямок своєї політики, прагнучи створити високий рівень кіберзахисту для своїх громадян, бізнесу та державних інституцій.

Нормативно-правове регулювання кібербезпеки в ЄС є багаторівневим та динамічним процесом. Його розвиток відображає прагнення країн-членів забезпечити як надійний захист критичної інфраструктури та персональних даних, так і сталий розвиток цифрової економіки. З цією метою ЄС прийняв низку важливих нормативно-правових актів, серед яких особливе місце займають Директива NIS та її оновлена версія – NIS2, GDPR, DORA, а також PCI DSS [1-4]. Кожен із цих документів має свою сферу застосування, особливості та завдання, однак усі вони спрямовані на досягнення спільної мети – забезпечення кіберстійкості та цифрової довіри в Європейському Союзі.

Аналіз та порівняння цих нормативних актів дозволяє глибше зрозуміти, якими шляхами ЄС намагається врегулювати питання кібербезпеки, ураховуючи постійні технологічні зміни та еволюцію кіберзагроз. Особливої актуальності це набуває в умовах геополітичних викликів, зокрема в контексті агресії російської федерації проти України, що ще більше підкреслило важливість кібербезпеки для захисту державних інституцій та критичних сервісів. Метою цієї статті є дослідження та порівняння основних нормативно-правових документів Європейського Союзу у сфері кібербезпеки, визначення їхніх спільних рис та відмінностей, а також аналіз впливу на держави-члени та бізнес.

2. Ключові нормативно-правові документи ЄС

Атаки на основі ШІ можна класифікувати за багатьма різними параметрами. Так, наприклад, кілька систем класифікації атак були представлені в роботах [7, 8]. Також NIST в своєму звіті щодо атак на машинне навчання ШІ представив різні типи класифікації [1, 25]. В цьому розділі будуть розглянуті деякі типи класифікацій.

2.1. Директива NIS2 та її особливості

CNIS2 Directive (Network and Information Security Directive 2) – це оновлена структура Європейського Союзу з кібербезпеки, яка замінює оригінальну Директиву NIS (2016) [1]. Вона розроблена для посилення кібербезпеки в Європейському Союзі шляхом встановлення високого загального рівня безпеки для мережевих та інформаційних систем. Її основною метою є

ISSN 2519-2310 (Online) Комп'ютерні науки та кібербезпека. 2025. № 1 (27)

посилення кіберстійкості критичних секторів, а також забезпечення уніфікованих стандартів кіберзахисту для всіх суб'єктів цифрової економіки, незалежно від їхньої галузі та розміру.

Директива ставить за мету кілька ключових завдань, які разом утворюють системний підхід до кібербезпеки. Насамперед це зміцнення можливостей держав-членів у сфері кібербезпеки, що передбачає розвиток національних стратегій, створення компетентних органів та ефективних механізмів реагування на інциденти. Другим важливим завданням є покращення співпраці між державами-членами, що сприяє своєчасному обміну інформацією про кіберзагрози та координації дій у випадку масштабних атак. Ще одна мета – зменшення фрагментації на внутрішньому ринку ЄС, адже різні підходи до регулювання в окремих країнах створювали прогалини у захисті цифрової інфраструктури. Нарешті, Директива запроваджує обов'язкові вимоги з управління ризиками та звітування про кіберінциденти, що суттєво посилює відповідальність організацій за стан кібербезпеки [1].

Основні вимоги Директиви охоплюють кілька напрямків. Перш за все, це управління ризиками, яке передбачає впровадження суб'єктами (операторами основних послуг та постачальниками цифрових послуг) технічних, операційних та організаційних заходів. До таких заходів належать [1]: систематичний аналіз ризиків та розробка політики безпеки; ефективне реагування на інциденти; забезпечення безперервності бізнесу та планування відновлення після інцидентів; безпека закупівель та використання ІТ-систем; а також політика моніторингу, тестування та аудитів кібербезпеки.

Другим важливим компонентом є звітування про інциденти. Відповідно до Директиви, суттєві інциденти повинні бути повідомлені до відповідних компетентних органів протягом 24 годин з моменту їх виявлення. При цьому організація має надати повне повідомлення про інцидент не пізніше, ніж через 72 години, а остаточний звіт – не пізніше, ніж через один місяць після події. Такий підхід дозволяє державним органам оперативно реагувати на інциденти, координувати дії між суб'єктами ринку та зменшувати потенційний вплив на суспільство [1].

Окрему увагу Директива приділяє постачальникам ІКТ-послуг, встановлюючи обов'язок оцінки та контролю ланцюга постачання. Це означає, що організації повинні впроваджувати заходи безпеки для забезпечення надійності своїх партнерів та контрагентів, а також враховувати кіберризики у договірних відносинах із критичними постачальниками. Такий підхід допомагає створити більш стійку екосистему цифрових послуг у межах ЄС [1].

Важливою особливістю NIS2 є посилення ролі керівних органів. Відтепер вони несуть відповідальність за схвалення та нагляд за впровадженням заходів кібербезпеки у своїх організаціях. У разі недотримання вимог Директиви керівники можуть бути притягнуті до відповідальності, що підкреслює важливість кібербезпеки як стратегічного питання для кожної компанії [1].

Зрештою, кожна держава-член ЄС зобов'язана призначити національні компетентні органи, які здійснюють нагляд за виконанням вимог Директиви. Це сприяє уніфікації підходів до моніторингу та контролю в межах ЄС, а також створює механізми ефективної взаємодії між країнами у сфері кібербезпеки.

2.2. Регламент GDPR та його особливості

General Data Protection Regulation (GDPR) – це Загальний регламент про захист даних, прийнятий Європейським Союзом і такий, що набув чинності 25 травня 2018 року [2]. Він є ключовим нормативним документом, який визначає правила обробки персональних даних громадян ЄС. Особливість цього документа полягає в його екстериторіальній дії: правила GDPR поширюються не лише на організації, що перебувають у межах Європейського Союзу, але й на будь-які компанії, що обробляють персональні дані громадян ЄС, незалежно від місця їхньої

ISSN 2519-2310 (Online) Computer Science and Cybersecurity. 2025. Issue 1 (27)

реєстрації [2]. У такий спосіб GDPR суттєво підвищує рівень захисту персональних даних і гарантує громадянам ЄС контроль над тим, як їхні дані збираються та використовуються.

Основними принципами, на яких базується GDPR, є законність, справедливість і прозорість обробки даних [2]. Це означає, що персональні дані повинні оброблятися на законних підставах, чесно та прозоро для суб'єкта даних. Наприклад, компанія має чітко повідомити людині, які дані збираються, для чого вони потрібні та як довго їх планують зберігати. Це створює довіру між користувачем і організацією та забезпечує право особи знати, що відбувається з її персональними даними.

Іншим ключовим принципом є обмеження мети: персональні дані повинні збиратися лише для конкретних, явних і законних цілей та не можуть оброблятися далі у спосіб, несумісний з цими цілями. Наприклад, якщо компанія збирає електронні адреси для реєстрації в сервісі, вона не має права використовувати ці дані для нав'язливої реклами без додаткової згоди користувача [2].

Мінімізація даних означає, що організації повинні збирати лише ті дані, які справді необхідні для досягнення задекларованих цілей. Таким чином, збір надлишкових даних забороняється, що запобігає зловживанням та підвищує безпеку особистої інформації. Принцип точності вимагає від організацій забезпечити актуальність даних: інформація повинна бути точною й оновлюватися за потреби [2]. Це важливо, щоб уникнути помилок або використання застарілих даних, що можуть призвести до порушення прав особи.

Принцип обмеження зберігання вказує на те, що персональні дані не повинні зберігатися довше, ніж це необхідно для досягнення цілей обробки [2]. Організації мають визначати строки зберігання даних і знищувати або анонімізувати їх після завершення обробки. Це суттєво знижує ризики витоку інформації та зловживання.

Цілісність і конфіденційність є одними з найважливіших вимог GDPR, оскільки вони вимагають забезпечити належну безпеку персональних даних, у тому числі їх захист від несанкціонованої або незаконної обробки, випадкової втрати, знищення чи пошкодження. Організації повинні впроваджувати технічні та організаційні заходи безпеки, такі як шифрування, регулярне резервне копіювання та контроль доступу, щоб гарантувати безпеку даних [2].

GDPR застосовується до організацій, розташованих у межах ЄС, які обробляють персональні дані. Крім того, його вимоги поширюються й на організації, що перебувають за межами ЄС, якщо вони пропонують товари чи послуги громадянам ЄС або моніторять поведінку осіб у межах ЄС [2]. Таким чином, навіть компанії, що не мають фізичної присутності в ЄС, підпадають під дію GDPR, якщо вони працюють із громадянами ЄС.

Порушення вимог GDPR може призвести до накладення значних штрафів, які становлять до 20 мільйонів євро або до 4% річного світового обороту компанії – залежно від того, яка сума є більшою [2]. Такі фінансові санкції стимулюють компанії дотримуватися принципів прозорої, законної та безпечної обробки персональних даних, а також створюють новий рівень відповідальності для бізнесу в цифровому середовищі.

2.3. Регламент DORA та його особливості

Digital Operational Resilience Act (DORA) – це Регламент Європейського Союзу, спрямований на посилення цифрової стійкості фінансового сектору. Він набув чинності 16 січня 2023 року, а його застосування розпочалось 17 січня 2025 року [3]. Цей регламент є ключовим кроком ЄС у забезпеченні стійкості фінансових установ до ризиків, пов'язаних із інформаційними та комунікаційними технологіями (ІКТ), зокрема кіберінцидентами та технічними збоями. DORA створює єдиний стандарт цифрової операційної стійкості для всіх

ISSN 2519-2310 (Online) Комп'ютерні науки та кібербезпека. 2025. № 1 (27)

фінансових організацій, щоб гарантувати їхню здатність витримувати, реагувати та відновлюватися після порушень у сфері ІКТ [3].

Однією з основних вимог DORA є управління ІКТ-ризиками, яке передбачає, що фінансові установи повинні впроваджувати комплексні рамки управління ризиками. Це включає ідентифікацію ризиків, їхню профілактику, своєчасне виявлення інцидентів, ефективне реагування та відновлення після подій [3]. Такий підхід дозволяє організаціям підготуватися до потенційних кіберзагроз та забезпечує безперервність надання фінансових послуг навіть під час кризових ситуацій.

Регламент також запроваджує обов'язкове звітування про інциденти: фінансові установи повинні повідомляти національні регуляторні органи про значні інциденти у сфері ІКТ. Це дозволяє державним органам оперативно реагувати на інциденти та координувати заходи для запобігання системним ризикам у фінансовому секторі [3]. Таким чином, DORA сприяє своєчасному обміну інформацією та знижує ризики поширення кіберзагроз.

Ще однією важливою вимогою є регулярне тестування цифрової операційної стійкості, яке передбачає проведення як базових, так і передових тестувань для певних установ [3]. Це допомагає визначити слабкі місця у системах та процесах організацій, а також забезпечити їхню готовність до реагування на кіберінциденти. Регулярні тренування та симуляції атак підвищують загальну кіберстійкість фінансової екосистеми ЄС.

Крім того, DORA передбачає управління ризиками третіх сторін. Фінансові установи зобов'язані встановлювати вимоги до ІКТ-постачальників, укладати відповідні контракти та забезпечувати належний моніторинг їхньої діяльності [3]. Це положення має на меті зменшити ризики, пов'язані із зовнішніми підрядниками, та гарантувати, що всі учасники цифрового ланцюга відповідають високим стандартам кібербезпеки.

Особливе значення в DORA приділяється обміну інформацією між фінансовими установами. Регламент заохочує обмін знаннями про кіберзагрози та вразливості, щоб підвищити колективну стійкість фінансової системи ЄС [3]. Такий підхід дозволяє швидко ідентифікувати нові загрози та ефективно реагувати на них у партнерстві з іншими учасниками ринку.

DORA поширюється на широкий спектр фінансових установ, серед яких: банки, страхові та перестрахові компанії, інвестиційні фірми, платіжні установи, електронні грошові установи, постачальники послуг з обміну криптовалют, центральні контрагенти, центральні депозитарії цінних паперів, управляючі компанії інвестиційних фондів, а також постачальники ІКТ-послуг, що обслуговують фінансові організації [3]. Така широка сфера застосування забезпечує комплексний підхід до захисту фінансової інфраструктури від кіберзагроз.

У разі порушення вимог DORA до фінансових установ можуть бути застосовані серйозні санкції. Це можуть бути адміністративні штрафи, розмір яких визначається національними регуляторами, а також обмеження або заборона на надання певних послуг [3]. Крім того, регулятори можуть вимагати усунення порушень у встановлені терміни. Такі заходи спрямовані на забезпечення високого рівня відповідальності фінансових установ та стимулювання їхнього дотримання єдиних правил цифрової операційної стійкості.

2.4. Атака отруєння даних

Атаки на етапі навчання ML називаються атаками отруєння [1, 9]. Під час атаки отруєння даних [5, 9] зловмисник контролює підмножину навчальних даних, вставляючи або змінюючи навчальні зразки. У атаці отруєння моделі [10] зловмисник контролює модель та її параметри. Атаки з отруєнням даних можуть застосовуватися до всіх парадигм навчання, тоді як атаки з отруєнням моделі є найбільш поширеними у федеративному навчанні [11], де клієнти

ISSN 2519-2310 (Online) Computer Science and Cybersecurity. 2025. Issue 1 (27)

надсилають локальні оновлення моделі на сервер, що обробляє вхідні дані, і в атаках на ланцюг поставок, де шкідливий код може бути доданий до моделі постачальниками технології моделі. Під федеративним навчанням тут мається на увазі метод машинного навчання, орієнтований на умови, в яких кілька суб'єктів (часто званих клієнтами) спільно навчають модель, при цьому дані, які використовуються для навчання, розподіляються децентралізовано. Це відрізняє його від машинного навчання, в якому дані зберігаються централізовано.

Перші атаки отруєння, виявлені в додатках кібербезпеки, були атаками на доступність проти генерації профілів хробака та класифікаторів спаму, які невибірково впливають на всю модель машинного навчання та, по суті, спричиняють атаку типу «відмова в обслуговуванні» для користувачів системи ШІ.

Атаки отруєння вважаються одними з найнебезпечніших серед атак на ШІ та можуть спричинити або порушення доступності, або порушення цілісності. Зокрема, атаки з порушенням доступності спричиняють деградацію моделі машинного навчання на всіх етапах, тоді як цільові та бекдорні атаки з отруєнням є більш прихованими та викликають порушення цілісності на невеликому наборі даних. Атаки отруєння використовують широкий спектр конкурентних можливостей, таких як отруєння даних, отруєння моделі, контроль міток, контроль вихідного коду та контроль тестових даних, що призводить до кількох підкатегорій атак отруєння. За моделлю загрози вони можуть використовуватись як у сценаріях білої скриньки, так і чорної скриньки [18], що були розглянуті в розділі 1.3 цієї статті.

Серед методів запобігання атакам отруєння даних виділяють два найефективніші [1]:

• Очищення навчальних даних. Ці методи використовують той факт, що отруєні набори зазвичай відрізняються від звичайних навчальних наборів, які не контролюються зловмисниками. Таким чином, методи очищення даних призначені для очищення навчального набору та видалення отруєних наборів перед виконанням навчання машинного навчання.

• Надійне навчання. Альтернативним підходом до пом'якшення атак з порушенням доступності є модифікація алгоритму навчання ML і проведення надійного навчання замість звичайного навчання. У кількох статтях визначено методи надійної оптимізації, такі як використання функції скорочених втрат [20] або випадкове згладжування для додавання шуму під час навчання [21].

2.5. Стандарти PCI DSS та їх особливості

Payment Card Industry Data Security Standard (PCI DSS) – це міжнародний стандарт безпеки, розроблений для захисту даних власників платіжних карток [4]. Його головною метою є забезпечення безпечного середовища для обробки, зберігання та передавання даних про кредитні картки. Всі організації, які приймають, обробляють або зберігають дані платіжних карток, повинні дотримуватися вимог цього стандарту, щоб зменшити ризик несанкціонованого доступу, шахрайства та витоку конфіденційної інформації.

PCI DSS є універсальним набором вимог, обов'язковим для компаній різного масштабу: від невеликих Інтернет-магазинів до великих фінансових установ. Він сприяє зміцненню кіберстійкості компаній, які працюють у сфері електронних платежів, та формує довіру споживачів до цифрових транзакцій. Стандарт включає дванадцять головних умов, які разом забезпечують комплексний захист платіжної інформації [4].

Першою вимогою є захист систем, що обробляють платіжні дані, включаючи фізичний захист обладнання та програмного забезпечення. Це передбачає використання фаєрволів, сегментації мережі та інших заходів для запобігання несанкціонованому доступу. Друга вимога стосується захисту даних власників карток від стороннього доступу, зокрема шляхом шифрування даних як під час передавання, так і під час зберігання [4].

ISSN 2519-2310 (Online) Комп'ютерні науки та кібербезпека. 2025. № 1 (27)

Також PCI DSS передбачає обов'язковий захист мереж, які обробляють дані про картки, від несанкціонованого доступу [4]. Це означає, що компанії повинні впроваджувати сучасні методи контролю доступу, моніторинг мережевого трафіку та системи виявлення вторгнень. Додатково стандарт вимагає захисту від шкідливого програмного забезпечення, яке може призвести до компрометації платіжних даних.

Суттєвою вимогою є обмеження доступу до платіжних даних: він має надаватися лише авторизованим користувачам, що мінімізує ризики внутрішніх загроз. PCI DSS також акцентує увагу на необхідності постійного тестування та розвитку безпеки систем, щоб своєчасно виявляти вразливості та запобігати можливим атакам [4].

Крім цього, компанії повинні впроваджувати та підтримувати діяльність ефективних політик безпеки, які встановлюють правила обробки платіжних даних та обов'язки працівників. Стандарт вимагає обмеження фізичного доступу до приміщень, де розміщене обладнання для обробки платіжних даних, а також постійного моніторингу інцидентів, які можуть свідчити про спроби несанкціонованого доступу [4].

PCI DSS наголошує на необхідності захисту від зовнішніх атак шляхом впровадження сучасних механізмів кіберзахисту, таких як системи виявлення вторгнень, міжмережеві екрани та антивірусні рішення [4]. Важливим є також управління підтримкою та обслуговуванням обладнання й програмного забезпечення, що передбачає регулярні оновлення систем та патч-менеджмент. Остання, але не менш важлива вимога – розробка та впровадження політик і процедур безпеки, які гарантують відповідність усім вимогам стандарту.

PCI DSS також запроваджує систему класифікації компаній за рівнями залежно від кількості оброблюваних транзакцій на рік [4]. Це дозволяє встановлювати різний рівень вимог до сертифікації залежно від обсягів діяльності. До першого рівня належать компанії, що обробляють понад 6 мільйонів транзакцій на рік, для яких передбачено найжорсткіші вимоги до безпеки. До другого рівня входять організації, що обробляють від 1 до 6 мільйонів транзакцій, а до третього – компанії з обсягом від 20 тисяч до 1 мільйона транзакцій на рік. Нарешті, до четвертого рівня належать компанії, які обробляють до 20 тисяч транзакцій щороку [4]. Такий підхід дозволяє гнучко адаптувати вимоги до різних бізнесів, одночасно забезпечуючи високий рівень захисту даних платіжних карток у всьому фінансовому ланцюгу.

3. Порівняння документів

3.1. Порівняння нормативно-правових документів ЄС

Для порівняльного аналізу нормативно-правових документів ЄС було обрано наступні критерії:

- мета регулювання;
- категорії суб'єктів регулювання;
- обов'язки організацій;
- вимоги до кібербезпеки;
- санкції за порушення;
- взаємозв'язки між документами.

Результати порівняння наведено у таблиці 1.

ISSN 2519-2310 (Online) Computer Science and Cybersecurity. 2025. Issue 1 (27)

Таблиця 1 – Порівняння нормативно-правових документів ЄС за критеріями
Table 1 – Comparison of EU Regulatory Documents Based on Criteria

Критерій Criteria	**GDPR**	**DORA**	**NIS2**	**PCI DSS**
Мета регулювання **The purpose of regulation**	Захист персональних даних фізичних осіб у межах ЄС Protection of personal data of individuals within the EU	Забезпечення цифрової стійкості фінансових установ до ІКТ-ризиків Ensuring the digital resilience of financial institutions to ICT risks	Підвищення рівня кібербезпеки мережевих та інформаційних систем в ЄС Enhancing the cybersecurity level of network and information systems in the EU	Забезпечення безпеки даних платіжних карток Ensuring the security of payment card data
Категорії суб'єктів **Categories of subjects**	Усі організації, що обробляють персональні дані громадян ЄС All organizations that process personal data of EU citizens	Фінансові установи та ІКТ-постачальники, що працюють з ними Financial institutions and ICT providers working with them	Суб'єкти з «високим ступенем критичності», включаючи цифрові, енергетичні, медичні, транспортні сектори Entities with a high degree of criticality, including the digital, energy, healthcare, and transport sectors	Організації, які обробляють, передають або зберігають платіжні карткові дані Organizations that process, transmit, or store payment card data
Обов'язки організацій **Responsibilities of organizations**	Законна обробка даних, обмеження мети, мінімізація, точність, захист Lawful data processing, purpose limitation, data minimization, accuracy, protection	Управління ІКТ-ризиками, інцидент-репортинг, тестування, управління ризиками третіх сторін ICT risk management, incident reporting, testing, third-party risk management	Управління ризиками, звітування про інциденти, оцінка постачальників, відповідальність керівництва Risk management, incident reporting, supplier assessment, management accountability	Забезпечення 12 основних умов безпеки, включаючи шифрування, контроль доступу, моніторинг Ensuring 12 core security requirements, including encryption, access control, and monitoring
Вимоги до кібербезпеки **Cybersecurity requirements**	Забезпечення конфіденційності, цілісності й доступності даних Ensuring the confidentiality, integrity, and availability of data	Впровадження ІКТ-безпеки: виявлення, захист, реагування, відновлення, аудит Implementation of ICT security: detection, protection, response, recovery, audit	Впровадження технічних, організаційних заходів, тестування, безперервність бізнесу Implementation of technical and organizational measures, testing, and business continuity	Технічні та організаційні заходи для захисту даних платіжних карт та систем, що їх обробляють Technical and organizational measures to protect payment card data and the systems that process them

ISSN 2519-2310 (Online) Комп'ютерні науки та кібербезпека. 2025. № 1 (27)

Санкції за порушення **Sanctions for violations**	До 20 млн. євро або до 4% світового обороту Up to €20 million or up to 4% of global turnover	Залежно від рішення національного регулятора: штрафи, обмеження діяльності, вимога усунення порушень Depending on the decision of the national regulator: fines, activity restrictions, requirement to eliminate violations	Адміністративні санкції, потенційна відповідальність керівників організації Administrative sanctions, potential liability of the organization's executives	Санкції від платіжних систем, відключення від обробки платежів, фінансові штрафи Sanctions from payment systems, disconnection from payment processing, financial penalties
Взаємозв'язки між документами **Interrelationships between documents**	GDPR фокусується на персональних даних, може доповнювати NIS2 або DORA GDPR focuses on personal data and can complement NIS2 or DORA	Може перетинатися з NIS2 (особливо в аспектах кібербезпеки) і з GDPR (якщо йдеться про персональні дані) May overlap with NIS2 (especially in cybersecurity aspects) and with GDPR (when personal data is involved)	Поширюється на ті самі галузі, що й DORA, та взаємодіє з GDPR у частині інцидентів і даних Applies to the same sectors as DORA and interacts with GDPR in terms of incidents and data	Може доповнювати GDPR щодо захисту фінансових персональних даних; має спільні цілі з DORA/NIS2 May complement GDPR in protecting financial personal data; shares common goals with DORA/NIS2

3.2. Порівняння нормативно-правових документів за впливом на бізнес та IT-інфраструктуру

GDPR в першу чергу впливає на організаційні процеси, змушуючи компанії переглянути політики конфіденційності, налагодити механізми для обробки запитів суб'єктів даних, наприклад: запити на доступ, виправлення чи видалення даних. Та призначити відповідального за захист даних. В IT-інфраструктурі це означає необхідність впровадження інструментів шифрування, анонімності, контролю доступу та логування. GDPR значно підвищує рівень захисту персональних даних, але майже не стосується фінансових транзакцій безпосередньо [2].

На відміну від GDPR, DORA зосереджується на цифровій стійкості фінансових організацій. Він впливає на бізнес через потребу в управлінні ІКТ-ризиками, регулярному тестуванні готовності до інцидентів, а також у створенні механізмів реагування та відновлення. Компанії повинні укладати чіткі контракти з IT-постачальниками, забезпечувати безперервність бізнесу навіть у разі кібератак або технічних збоїв. З технічного боку, це означає впровадження систем моніторингу, автоматизованих тестів безпеки, резервного копіювання, сегментації мереж тощо. DORA опосередковано посилює захист і персональних, і фінансових даних завдяки покращенню стійкості систем [2, 3].

Директива NIS2 також має значний вплив на організаційну структуру компаній, особливо тих, що належать до критично важливих секторів. Вона вимагає системного підходу до управління кіберризиками, створення інцидент-менеджменту, безпеки ланцюгів постачання, а також прямої відповідальності керівництва за впровадження кіберзахисту. В

ISSN 2519-2310 (Online) Computer Science and Cybersecurity. 2025. Issue 1 (27)

IT-інфраструктурі це призводить до запровадження засобів моніторингу, журналювання, тестування безпеки, а також процедур забезпечення безперервності бізнесу. Як результат, посилюється захист як персональних, так і фінансових даних у рамках загальної кібербезпеки [1].

PCI DSS безпосередньо впливає на компанії, які працюють із платіжними картками. Він вимагає стандартизації процедур обробки платіжних даних, впровадження політик безпеки, обмеження доступу, фізичного захисту обладнання, а також регулярного тестування безпеки. В IT-сфері компанії повинні забезпечити шифрування даних карток, встановлення брандмауерів, систем виявлення вторгнень, обмеження доступу до систем. PCI DSS безпосередньо націлений на захист фінансових транзакцій та даних власників карток, забезпечуючи надійність платіжної інфраструктури [1].

4. Міжнародні аналоги нормативно-правових актів ЄС

Регуляторна система Європейського Союзу у сфері цифрової безпеки вважається однією з найрозвиненіших у світі. Водночас важливо зазначити, що багато інших країн та міжнародних організацій створили власні нормативні документи, які за своєю метою, структурою та принципами є подібними до європейських.

Зокрема, регламент GDPR, що встановлює єдині правила захисту персональних даних у межах ЄС, має численні відповідники у законодавствах інших країн. У США діють кілька галузевих актів, таких як HIPAA – закон про захист персональних медичних даних, який містить вимоги щодо безпеки інформаційних систем у сфері охорони здоров'я [5]. Для освітнього сектору розроблено FERPA, що гарантує студентам і їхнім батькам право на доступ до освітніх записів та їхнє коригування [6]. Водночас в окремих штатах запроваджуються ширші закони про конфіденційність, зокрема CCPA у Каліфорнії, що дає споживачам право знати, які персональні дані збираються про них, і вимагати їх видалення [7]. У Південній Америці Бразилія прийняла закон LGPD, який вважається майже повним аналогом GDPR – як за принципами обробки даних, так і за правами користувачів [8]. Окрему роль відіграє Конвенція № 108 Ради Європи, до якої приєдналася і Україна, – вона є першою міжнародною угодою, що встановила обов'язкові правила щодо захисту персональних даних як у державному, так і приватному секторах [9].

Регламент DORA, спрямований на забезпечення цифрової стійкості фінансового сектору ЄС, також має функціональні аналоги. У США Міжбанківська координаційна рада розробила FFIEC Cybersecurity Assessment Tool – інструмент для оцінювання рівня кіберзахищеності фінансових установ [10]. Структурно близькою до DORA є і модель NIST Cybersecurity Framework, яка охоплює повний життєвий цикл безпеки: від ідентифікації загроз до відновлення після інцидентів [11].

Директива NIS2, яка зобов'язує операторів критичної інфраструктури впроваджувати кіберзахист, має численні відображення у регулюваннях інших країн і сфер. Найбільш очевидною паралеллю є NIST Cybersecurity Framework у США, який є універсальним інструментом для підвищення кіберстійкості організацій [10]. У галузі енергетики в США та Канаді застосовуються стандарти NERC CIP – обов'язкові норми безпеки для енергетичних компаній [12]. Велика Британія після виходу з ЄС ухвалила UK NIS Regulations – документ, що фактично є національною адаптацією європейської директиви NIS [13]. До найближчих до NIS2 за ініціативами документів також можна віднести міжнародний стандарт ISO/IEC 27001, який стосуються управління інформаційною безпекою та забезпечення безперервності діяльності [14].

У сфері платіжних систем міжнародний стандарт PCI DSS знаходить чіткі відповідники у низці документів. Зокрема, ISO/IEC 27001 як базовий стандарт управління інформаційною

ISSN 2519-2310 (Online) Комп'ютерні науки та кібербезпека. 2025. № 1 (27)

безпекою часто використовується у фінансовому секторі [14]. У США подібну функцію виконують рекомендації NIST SP 800-53 та 800-171 – це комплексні каталоги технічних, фізичних і адміністративних заходів захисту інформаційних систем [15, 16]. У галузі електронної комерції та хмарних послуг активно використовується стандарт SOC 2, що дозволяє оцінити рівень безпеки, конфіденційності та моніторингу в компаніях, які обробляють дані клієнтів [17].

Загалом, міжнародна практика демонструє, що ключові принципи цифрової безпеки – управління ризиками, прозорість, захист персональних даних, відповідальність організацій – стали універсальними.

5. Висновки

1. У дослідженні проаналізовано основні нормативно-правові документи ЄС у сфері кібербезпеки: Директиву NIS2, Регламент GDPR, Регламент DORA та стандарт PCI DSS.

2. Усі ці документи спрямовані на посилення кіберстійкості та цифрової довіри в ЄС, але мають різні акценти:

- NIS2 фокусується на безпеці мережевих та інформаційних систем, зокрема для критичної інфраструктури.
- GDPR створює систему захисту персональних даних, встановлює права громадян і правила обробки інформації.
- DORA регулює цифрову стійкість фінансового сектору, зокрема управління ІКТ-ризиками.
- PCI DSS встановлює вимоги до безпеки обробки платіжної інформації, особливо для e-commerce компаній.

3. Також документи мають багато точок перетину:

- GDPR доповнює NIS2 і DORA у питаннях захисту персональних даних під час кіберінцидентів.
- DORA і NIS2 мають спільні вимоги щодо управління ризиками та кіберстійкості.
- PCI DSS виступає як практичний інструмент реалізації вимог з безпеки платіжних даних, що підкріплює цифрову безпеку загалом.

4. Ефективне впровадження цих документів у країнах-членах ЄС сприятиме: підвищенню кіберстійкості; зміцненню довіри до цифрових технологій; сталому розвитку цифрової економіки; закріпленню позицій ЄС як глобального лідера у сфері цифрової безпеки.

Конфлікт інтересів

Автори повідомляють про відсутність конфлікту інтересів.

References

1. Directive (EU) 2022/2555 of the European Parliament and of the Council (2022) On measures for a high common level of cybersecurity across the Union (NIS 2 Directive). Official Journal of the European Union. – URL: https://surl.li/xpktzi
2. Regulation (EU) 2016/679 of the European Parliament and of the Council (2016) On the protection of natural persons with regard to the processing of personal data (GDPR). Official Journal of the European Union. – URL: https://surl.lt/zdilhe
3. Regulation (EU) 2022/2554 of the European Parliament and of the Council (2022) On digital operational resilience for the financial sector (DORA). Official Journal of the European Union. – URL: https://surl.li/lvfsva

4. PCI DSS Certification (n.d.) Payment Card Industry Data Security Standard. PCI Security Standards Council. – URL: https://getpci.com/
5. Centers for Medicare & Medicaid Services (n.d.) HIPAA – Health Insurance Portability and Accountability Act. – URL: https://www.cms.gov/priorities/key-initiatives/burden-reduction/administrative-simplification/hipaa
6. U.S. Department of Education (n.d.) What is FERPA? – URL: https://surl.lu/oizpnq
7. Office of the Attorney General of California (n.d.) California Consumer Privacy Act (CCPA). – URL: https://oag.ca.gov/privacy/ccpa
8. Brazilian Government (n.d.) General Personal Data Protection Act (LGPD). – URL: https://lgpd-brazil.info/
9. Council of Europe (1981) Convention No. 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data. – URL: https://ippi.org.ua/vid-redaktsiinoi-kolegii-konventsiya-%E2%84%96-108-radi-%D1%94vropi-%E2%80%9Cpro-zakhist-osib-u-zv%E2%80%99yazku-z-avtomatizovano
10. Federal Financial Institutions Examination Council (FFIEC) (n.d.) Cybersecurity Assessment Tool. – URL: https://www.ffiec.gov/resources/cat
11. Federal Trade Commission (n.d.) What is the NIST Cybersecurity Framework? – URL: https://www.ftc.gov/business-guidance/small-businesses/cybersecurity/nist-framework
12. North American Electric Reliability Corporation (n.d.) Critical Infrastructure Protection (NERC CIP). – URL: https://www.techtarget.com/searchsecurity/definition/North-American-Electric-Reliability-Corporation-Critical-Infrastructure-Protection-NERC-CIP
13. UK Government (2018) The Network and Information Systems Regulations 2018 (NIS Regulations 2018). – URL: https://surl.li/zahavj
14. ISO/IEC (2022) ISO/IEC 27001:2022 – Information Security, Cybersecurity and Privacy Protection. – URL: https://www.iso.org/standard/27001
15. National Institute of Standards and Technology (2020) NIST Special Publication 800-53 Rev. 5: Security and Privacy Controls for Information Systems and Organizations. – URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r5.pdf
16. National Institute of Standards and Technology (2020) NIST Special Publication 800-171 Rev. 2: Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations. – URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-171r2.pdf
17. Palo Alto Networks (n.d.) What is SOC 2 Compliance? – URL: https://www.paloaltonetworks.com/cyberpedia/soc-2