

Volodymyr Mishchenko

*Doctor of Economic Sciences, Professor,
Head of the Sector of Digital Economy
Institute for Economics and Forecasting of the NAS of Ukraine*

Svitlana Naumenkova

*Doctor of Economic Sciences, Professor,
Professor of the Department of Finance
Taras Shevchenko National University of Kyiv*

Svitlana Mishchenko

*Doctor of Economic Sciences, Professor,
Professor of the Department of Finance
State University of Trade and Economics*

**MANAGEMENT OF ARTIFICIAL INTELLIGENCE
UTILIZING VALUE-BASED PRINCIPLES**

Summary

The constant expansion of the fields of application of artificial intelligence technologies requires deliberate support to enhance their positive effects on the economy and society while reducing potential risks and negative consequences through the development and practical implementation of the principles of ethical and responsible management. Based on a comparative analysis, the experience of international organizations and leading countries of the world regarding the development and practical implementation of the principles of ethical, reliable, safe, and socially responsible design, implementation and use of technologies and systems of artificial intelligence in various sectors of the economy and spheres of activity are summarized. The content of the concept of artificial intelligence is clarified and the main tasks of developing unified principles and management methods based on the use of internationally recognized standards are defined. The peculiarities of the processes of development and introduction of principles and standards in different countries of the world are characterized and their periodization is carried out. The main types of organizational and management structures that manage the implementation, use and ensuring international compatibility of artificial intelligence technologies and systems are classified and characterized. The main approaches of international organizations and governments of individual countries regarding the development of legislative and regulatory acts in the field of regulation and control over the use of artificial intelligence are highlighted. The task of regulatory and control bodies to ensure the reliable use of artificial intelligence and reduce risks is defined. Recommendations have

been developed for Ukraine regarding the management of artificial intelligence based on the use of value-based principles, the use of which will contribute to strengthening the digital sovereignty of the country, accelerating the recovery of the economy, and increasing the level of international competitiveness.

Вступ

Технології штучного інтелекту все активніше використовуються в різних галузях економіки та сферах діяльності людей у всьому світі. Революційною подією у цій галузі стала поява технологій нового покоління – генеративного штучного інтелекту, який сприяв появі нових джерел створення вартості, суттєвому підвищенню продуктивності праці та зниженню транзакційних витрат, що трансформувало існуючі методи і підходи до регулювання та управління впровадженням і використанням інформаційно-комунікаційних технологій.

Технології штучного інтелекту прискорили розвиток ринків даних і механізмів роботи з ними, машинного та глибокого навчання, мовних технологій, стимулювали значні зміни в організації наукових досліджень, охороні здоров'я та медицині, освіті, підготовці кадрів та трудових відносинах загалом, сприяли захисту прав і свобод людини, зменшенню економічної та соціальної нерівності, забезпеченню екологічної стійкості, розширенню можливостей та підвищенню творчих здібностей людей, соціальному діалогу та залученню недостатньо представлених груп населення, посилили міждисциплінарне та міжнародне співробітництво.

Однак, разом із новими можливостями для держави, бізнесу та суспільства виникли й певні загрози та ризики щодо упередженого, необґрунтованого, неправомірного, неетичного або дискримінаційного використання технологій та систем штучного інтелекту, а в окремих випадках країни можуть зіткнутися із вимушеною втратою свого цифрового суверенітету. Сьогодні майже всі країни наражаються на такі ризики, а тому практично однотайною є думка багатьох науковців і експертів щодо необхідності розроблення та запровадження єдиних принципів, підходів і стандартів щодо впровадження надійних, етичних і соціально відповідальних систем штучного інтелекту.

Така ситуація висуває на перший план проблеми, пов'язані з розробленням єдиних технічних і технологічних стандартів, а також організацією дієвих та ефективних систем управління процесами регулювання та контролю за проектуванням, впровадженням і використанням систем штучного інтелекту для прискорення його використання та запобігання потенційним ризикам.

Головними завданнями міжнародних організацій та урядів окремих країн щодо розроблення єдиних принципів і методів управління штучним інтелектом є формування відповідної законодавчої та нормативно-

правової бази, яка б могла забезпечити прозорість їх функціонування впродовж усього життєвого циклу на основі повного розкриття інформації та належної підзвітності; надання інформації про фактори та процеси прийняття рішень і можливості відшкодування збитків від необґрунтованих або небезпечних рекомендацій; зменшення упередженості результатів, прогнозів і рішень, які надають системи штучного інтелекту; створення незалежних органів нагляду для аудиту використання алгоритмів та інші, що й обумовлює актуальність дослідження.

Розділ 1. Методологічні та методичні засади управління штучним інтелектом

У зв'язку зі складністю та багатоаспектністю поняття «штучний інтелект» у науковій літературі ще не існує його єдиного визначення. В різних джерелах штучний інтелект характеризують як технологію, інструмент, імітаційну модель, що може «навчатися», людино-машинну систему, послідовність алгоритмів тощо, або визначають через ключові характеристики, функції, призначення чи потенційні можливості [1].

Наприклад, фахівці Організації економічного співробітництва та розвитку (ОЕСР) розглядають штучний інтелект як «цифрову технологію загального призначення, використання якої покликане сприяти підвищенню рівня життя та добробуту людей, забезпеченню позитивної та стійкої глобальної економічної діяльності, розширенню можливостей для впровадження інновацій та підвищенню продуктивності праці, а також вирішенню ключових глобальних проблем» [2]. Європейський парламент визначає штучний інтелект як інструмент, що використовується програмою для відтворення поведінки, пов'язаної з людиною, такої як міркування, планування та творчість [3]. У «Концепції розвитку штучного інтелекту в Україні» це поняття характеризується як організована сукупність інформаційних технологій, із застосуванням яких можливо виконувати складні комплексні завдання шляхом використання системи наукових методів досліджень і алгоритмів обробки інформації, отриманої або самостійно створеної під час роботи, а також створювати та використовувати власні бази знань, моделі прийняття рішень, алгоритми роботи з інформацією та визначати способи досягнення поставлених завдань [4; 5].

На наш погляд, сьогодні штучний інтелект можна розглядати як комплексну системну сукупність технічних і технологічних рішень, розроблених на основі використання відповідних логічних алгоритмів, сформованих шляхом виявлення закономірностей функціонування окремих процесів або явищ, що дають можливість імітувати здатність людини до логічного мислення та виробляти на цій основі певні

узагальнення, рішення або рекомендації. Однак, з часом, враховуючи високі темпи розвитку сучасних технологій, зміст цього поняття, можливості та функції штучного інтелекту будуть змінюватися.

Складність проєктування та використання технологій і систем штучного інтелекту обумовлює необхідність забезпечення їх регулювання та управління на основі запровадження єдиних принципів і підходів, що ґрунтуються на використанні міжнародно визнаних стандартів. З метою вирішення цих завдань ОЕСР в 2019 р. розробила «Принципи штучного інтелекту», які стали глобальним стандартом для впровадження та використання цієї технології [2]. У 2021 р. ЮНЕСКО прийняла «Рекомендації з етики штучного інтелекту», в яких визначено загальні принципи та цінності, якими повинні керуватися розробники та користувачі при впровадженні та використанні технологій і систем штучного інтелекту з метою забезпечення стійкого та безпечного розвитку, захисту прав і свобод людини, а також протидії зміні клімату та соціальним катаклізмам [6].

У 2023 р. лідери країн G7 створили міністерський форум «Хіросімський процес штучного інтелекту», завдання якого полягає у підтримці сучасних систем штучного інтелекту та розробці міжнародних стандартів його надійності. Уряди країн G7 прийняли «Заяву про генеративний штучний інтелект», у якій визначено проблеми, пов'язані із захистом даних і конфіденційністю [7]. Крім того, міністри цифрових технологій країн G7 опублікували заяву, в якій проведено ґрунтовний аналіз проблем і можливостей використання штучного інтелекту, оголошено про розроблення міжнародних керівних принципів і кодексу поведінки для організацій, що проєктують системи штучного інтелекту [8].

Практична важливість питань використання нової технології обумовила посилення уваги урядів та бізнесу до управління штучним інтелектом. Першим етапом запровадження принципів і стандартів використання технологій штучного інтелекту в багатьох країнах світу стало розроблення національних стратегій штучного інтелекту. Лідерами були Канада, Фінляндія та Японія, а в 2020 р. власну стратегію розробила й Україна [4]. В окремих країнах національні стратегії штучного інтелекту є частиною загальної стратегії цифрового розвитку, як, наприклад, «Цифрова Швейцарія». Найбільш ґрунтовні стратегії розроблені урядами Великобританії, Ізраїлю, Індії, Канади, Японії, США, Південної Кореї та Сінгапуру. Вони передбачають формування ефективних екосистем відповідального штучного інтелекту, які покликані сприяти вирішенню ключових економічних і соціальних проблем шляхом підтримки інклюзивного зростання, сталого розвитку,

реалізації людиноцентричних цінностей, справедливості та розширення соціальних можливостей і творчих здібностей людей.

Разом з тим, кожна стратегія має свої особливості, які стосуються важливих аспектів державної політики в галузі штучного інтелекту. Наприклад, відповідно до указу Президента США «Про Національну ініціативу в галузі штучного інтелекту», головною метою національної стратегії США є захист економічної та національної безпеки, громадянських свобод, американських цінностей та інтересів. У стратегії виділено п'ять головних напрямів діяльності: розширення досліджень і розробок, надання додаткових матеріальних і фінансових ресурсів, розроблення та впровадження стандартів, що стимулюють розробку «надійних, безпечних і сумісних систем штучного інтелекту», підготовка кадрів, підвищення рівня цифрової грамотності населення, а також розширення міжнародного співробітництва [9].

Головними завданнями японської «Стратегії штучного інтелекту 2022» є підвищення конкурентоспроможності національної промисловості та пом'якшення соціальних наслідків використання технологій штучного інтелекту в разі виникнення непередбачуваних обставин [10]. У стратегії Великобританії основна увага приділяється розробленню засад і механізмів національного та міжнародного управління системами штучного інтелекту з метою стимулювання інновацій та інвестицій, а також захисту суспільних і фундаментальних цінностей. Індія планує стати глобальним хабом у галузі штучного інтелекту для країн, що розвиваються, а також активно розробляє мовні технології [11]. У стратегії України головна увага приділяється питанням використання потенціалу штучного інтелекту для забезпечення цифрового суверенітету та реалізації національних пріоритетів у сфері оборони, безпеки, освіти, державного управління, а також у промисловості та сільському господарстві [4; 5; 12, с. 68; 13, с. 108].

З метою ефективної реалізації головних положень своїх національних стратегій щодо впровадження та використання технологій і систем штучного інтелекту уряди створюють відповідні організаційно-управлінські структури, серед яких можна виокремити чотири головні їх види.

1) Окремий урядовий орган, на який покладено обов'язки організації та контролю за реалізацією національної стратегії, а також регулювання питань, пов'язаних із впровадженням і використанням технологій штучного інтелекту (Європейська рада з питань штучного інтелекту; Національний офіс ініціатив зі штучного інтелекту в США (NAIPO); Урядовий офіс Великобританії з питань штучного інтелекту у складі Департаменту науки, інновацій та технологій; Бюро політики у галузі штучного інтелекту Південної Кореї).

2) Багатосторонні комітети, до складу яких входить широке коло зацікавлених сторін, у тому числі представники бізнесу, експерти та науковці. Завдання таких комітетів полягає у сприянні розвитку цифрових екосистем, нагляді за розробкою та впровадженням національної стратегії штучного інтелекту, а також у проведенні моніторингу її виконання (прикладом можуть бути Національний центр з питань штучного інтелекту Австралії; Національний офіс з просування штучного інтелекту нового покоління Китаю; Управління штучного інтелекту Саудівської Аравії; Національна рада зі штучного інтелекту Єгипту; Комітет з управління стратегією штучного інтелекту Бразилії; Рада з питань штучного інтелекту Сербії).

3) Окремі незалежні органи виконавчої влади зі спеціальним статусом, які відповідають за певні напрями діяльності або використання цифрових технологій (наприклад, Канадський інститут комісара зі штучного інтелекту і даних, Національне управління даних Китаю) [11]. В Україні, відповідно до законопроекту № 6177, передбачено створення Національної комісії з питань захисту персональних даних та доступу до публічної інформації, головним завданням якої повинні бути імплементація міжнародних стандартів у сфері захисту права на доступ до інформації та захист персональних даних [14].

4) Експертні консультативні комітети та наглядові групи, створені з метою розроблення рекомендацій, підтримки ініціатив щодо розвитку штучного інтелекту та цифрової етики даних, вдосконалення нормативно-правової бази, консультування уряду з питань регулювання впровадження та використання систем штучного інтелекту, оцінки потенційних ризиків тощо (Шведський комітет у справах технологічних інновацій та етики; Ізраїльський урядовий центр з регулювання штучного інтелекту; Форум зацікавлених сторін Південної Кореї; канадська Консультативна рада з питань штучного інтелекту; Консультативна рада з етичного використання штучного інтелекту і даних Сінгапуру; Міжвідомча робоча група з питань штучного інтелекту Швейцарії; Національний консультативний комітет Франції). Наприклад, прем'єр-міністр Франції звернувся до Комітету з проханням надати висновки з етичних питань, пов'язаних з цифровими додатками, що використовують машинне навчання, для розмовних агентів автономних автомобілів і сфери медичної діагностики та здоров'я [11; 16, с. 286; 17, с. 12].

Крім того, з метою оцінки стану впровадження стратегій штучного інтелекту, посилення моніторингу їх функціонування, а також співпраці із зацікавленими сторонами та обміну знаннями ОЕСР створила Обсерваторію штучного інтелекту (ОЕСР. AI Policy Observatory). Аналогічні обсерваторії створено у Бразилії, Німеччині, Чилі, Чехії, Франції та у канадській провінції Квебек [11].

Необхідність забезпечення надійного використання та зниження рівня ризиків, пов'язаних із розробкою, впровадженням і використанням технологій і систем штучного інтелекту, обумовлює потребу в організації адекватної системи управління, регулювання та контролю. На думку фахівців ОЕСР, для цього необхідно: 1) визначити чіткі принципи впровадження та використання систем штучного інтелекту; 2) розробити відповідну нормативно-правову базу; 3) розробити та запровадити міжнародні стандарти; 4) сформувати контрольоване середовище для функціонування та розвитку екосистем штучного інтелекту [11].

У 2023 р. 17 країн світу розробили національні принципи та етичні правила проєктування, впровадження і використання систем штучного інтелекту, які в більшості випадків співпадають з принципами, проголошеними ОЕСР та ЮНЕСКО. Так, уряд Південної Кореї розробив «Керівні принципи етики штучного інтелекту», які включають десять вимог: права людини, захист конфіденційності, різноманіття, попередження шкоди, суспільний добробут, солідарність, управління даними, підзвітність, безпека та прозорість. Схожі документи розроблено в Австралії, Аргентині, Великобританії, Ізраїлі, Китаї, ОАЕ, Сінгапурі, США, Таїланді, Швейцарії, Японії та інших країнах. Сінгапурський звіт знань з етики та управління технологіями штучного інтелекту призначений для розв'язання практичних проблем, що виникають у зв'язку з необхідністю підтримки безпеки людей та реалізації людиноцентричних цінностей. В Індії розроблено Принципи управління відповідальними системами штучного інтелекту [11].

Найбільших результатів, зокрема й у сфері імплементації принципів ОЕСР, досягла Канада, яка є одним із лідерів у галузі штучного інтелекту. У Канаді функціонує 20 державних дослідницьких лабораторій, 75 інкубаторів та акселераторів, а також понад 850 стартапів, діяльність яких пов'язана з системами штучного інтелекту.

З метою врегулювання питань нагляду та контролю за системами штучного інтелекту Канада розробила Цифрову Хартію, яка включає Закон про штучний інтелект і дані (AIDA), Закон про захист конфіденційності споживачів і Закон про трибунал для захисту особистої інформації та даних. Для захисту конфіденційності споживачів у Канаді розроблено правила використання автоматизованих систем у сфері прийняття управлінських рішень і надання державних послуг, запроваджується регулювання технологій штучного інтелекту, які мають системне значення, а також сформовано механізми відповідального використання систем штучного інтелекту, які захищають права людини [18; 19].

Європейський Союз вперше у світі реалізував комплексний підхід до формування механізмів регулювання штучного інтелекту, які орієнтовані

на людиноцентричні цінності, та розробив законопроект «Про штучний інтелект». У зазначеному законі передбачено механізми регулювання великих та високоризикових систем штучного інтелекту, які можуть наражати економіку та суспільство на системні ризики, введено визначення поняття «базова модель штучного інтелекту» (наприклад, генеративного), а також систем загального призначення (GPAI), які використовуються для досягнення кількох цілей. Згідно із законопроектом, заборонено використання штучного інтелекту в окремих сферах діяльності та життєзабезпечення, які потенційно можуть становити небезпеку для здоров'я людей, сприяти маніпулюванню їх поведінкою, впливати на результати виборів або іншим чином порушувати фундаментальні права. Зокрема, це стосується систем біометричної категоризації, які ґрунтуються на використанні таких характеристик як політичні, релігійні або філософські переконання, раса, стать, вік, особливі потреби; копіювання зображень обличчя для створення баз даних розпізнавання людей; соціального скорингу на основі поведінки тощо. Обмежено сфери використання штучного інтелекту правоохоронними органами, а також закріплено право громадян отримувати пояснення щодо рішень, запропонованих штучним інтелектом, та оскаржувати результати, якщо вони порушують права людей.

Важливе значення для практичного використання регуляторних механізмів має запропонована в законі класифікація ризиків систем штучного інтелекту залежно від рівня їх надійності та сфер застосування: неприйнятні; високі; обмежені; мінімальні (або відсутні). Крім того, закон передбачає, що на компанії, які порушують його положення, можуть бути накладені штрафи в розмірі від 35 млн. євро або 7% від світового обороту до 7,5 млн. євро чи 1,5% від обороту компанії. За більш суттєві порушення, наприклад, за використання штучного інтелекту для підсвідомого маніпулювання або вразливості людей, штрафні санкції можуть бути ще більшими. При цьому важливо зазначити той факт, що санкції можуть бути встановлені для всіх виробників технологій, операторів і користувачів систем штучного інтелекту, в тому числі й за межами ЄС [20].

Політику США в галузі регулювання штучного інтелекту визначають три головні документи: заява офісу президента «Збереження американського лідерства в галузі штучного інтелекту», «План Білля про права штучного інтелекту» та «Білл про права штучного інтелекту», в яких сформульовано принципи етичного, безпечного та відповідального розроблення, впровадження і використання технологій штучного інтелекту. Ці принципи, серед іншого, передбачають необхідність захисту конфіденційності та забезпечення контролю за доступом до даних

шляхом використання вбудованих у продукти та послуги засобів захисту, а також право користувачів відмовитись від обслуговування з використанням автоматизованих систем [21; 22; 23].

Управління ризиками функціонування систем штучного інтелекту в США здійснюється відповідно до настанов Національного інституту стандартів і технологій шляхом включення відповідних етичних принципів у технічні регламенти та керівництва, а головними напрямками цього процесу визначено: створення безпечних і відповідальних систем штучного інтелекту, запровадження алгоритмічного захисту від дискримінації, захист конфіденційності даних, обов'язковість чітких пояснень та інформування користувачів, створення можливостей для контролю з боку людини за рішеннями та рекомендаціями штучного інтелекту, а також визначення варіантів пом'якшення ризиків для прав громадян і демократичних цінностей [24].

У 2022 р. в Ізраїлі опубліковано «Принципи політики, регулювання та етики в галузі штучного інтелекту», в яких викладено правові норми та рекомендації у галузі регуляторної політики. Зокрема, запропоновано галузеве регулювання систем штучного інтелекту, що ґрунтуються на загальних підходах до оцінки та управління ризиками. При цьому пріоритет надається використанню «м'яких» методів на основі використання етичних принципів, стандартів і добровільного прийняття рекомендацій [25]. Уряд Великобританії підготував стратегічний документ «Регулювання штучного інтелекту: інноваційний підхід», у якому передбачено регулювання на основі використання галузевого підходу. Відповідно до документа, перевага надається добровільним заходам, тобто реалізується децентралізований підхід, який передбачає високий рівень гнучкості [1]. В Японії розроблено «Соціальні принципи людиноцентричного штучного інтелекту», до яких віднесено гідність, різноманіття, стійкість, надійність, підзвітність, прозорість, конфіденційність, безпеку [26].

Специфічною є практика регулювання процесів проектування, впровадження та використання систем штучного інтелекту в Китаї, який розробляє загальне законодавство щодо регулювання та нагляду за ними, поєднуючи добровільні етичні норми, галузеві стандарти та рекомендації з жорсткими адміністративними вимогами, згідно з якими запроваджуються обов'язкові горизонтальні правила. Крім того, в Китаї розроблено стандарти захисту даних для автоматизованого водіння, а також нормативні акти з алгоритмічного управління, згідно з якими відповідальність за рекомендації, засновані на алгоритмах, покладається на постачальників цифрової інформації [27].

Узагальнюючи викладене, можна зробити висновок про те, що сьогодні у галузі розроблення законодавства та нормативних актів щодо

регулювання та контролю за використанням технологій і систем штучного інтелекту країни застосовують два основні підходи. Перший – вертикальний, який передбачає уніфікацію законодавства і ґрунтується на використанні єдиної системи правил і вимог для всіх учасників екосистеми, галузей економіки і сфер діяльності, гнучко поєднуючи загальні правила із галузевими стандартами та технічними умовами. Найбільш чітко такий підхід прослідковується в ЄС, Канаді, Бразилії та Китаї.

Другий підхід – горизонтальний (галузевий), згідно з яким встановлюються міжсекторальні етичні принципи, які є добровільними для використання, а галузеві регулятори визначають конкретні правила та умови впровадження і використання систем штучного інтелекту в окремих секторах. Така практика набула поширення в більшості країн світу, зокрема, у Великобританії, Ізраїлі, Індії, ОАЕ, Саудівській Аравії, Сінгапурі, США та Японії [11; 28, с. 42].

Розділ 2. Практичні аспекти реалізації ціннісних принципів управління штучним інтелектом

Відповідно до керівних настанов ОЕСР та ЮНЕСКО, головними ціннісними принципами, якими повинні керуватися розробники і користувачі технологій і систем штучного інтелекту, є такі.

1. Принцип «Інклюзивне зростання, стійкий розвиток і добробут» передбачає, що відповідальне управління технологіями та системами штучного інтелекту повинно сприяти підвищенню продуктивності праці, інклюзивному зростанню, стійкому розвитку та добробуту всіх членів суспільства; зменшенню економічної, соціальної, гендерної та інших видів нерівності; забезпеченню ефективного міждисциплінарного співробітництва зацікавлених сторін; підтримці соціального діалогу з громадянами та залученню в економічне і громадське життя недостатньо представлених груп населення; використанню технологій штучного інтелекту для підтримки екологічної стійкості та захисту навколишнього середовища; розширенню можливостей та розвитку творчих здібностей людей [2; 6; 29, с. 187; 30, с. 65].

З метою реалізації цих завдань діяльність урядів повинна бути зосереджена на реалізації таких напрямів політики, які сприяють зменшенню нерівності, розширенню співробітництва за участі зацікавлених сторін, посиленню соціального діалогу, використанню штучного інтелекту для екологічної безпеки та розширення творчого потенціалу людей. Найбільших успіхів у цьому напрямі досягли уряди Канади, Німеччини, США, Франції та країн ЄС (табл. 1).

Таблиця 1

**Завдання політики та приклади реалізації вимог
принципу «Інклюзивне зростання, стійкий розвиток і добробут»
в окремих країнах світу**

Завдання політики	Приклади реалізації
Зменшення економічної, соціальної та гендерної нерівності	<ul style="list-style-type: none"> • Французька програма «IA Booster» для зменшення економічної та соціальної нерівності в діяльності малих і середніх підприємств (2021) • Ініціатива Інституту Алана Тюрінга «Жінки у науці про дані та штучному інтелекті» (2023)
Міждисциплінарне співробітництво за участі зацікавлених сторін	<ul style="list-style-type: none"> • Німецька ініціатива «Громадянське кодування – інноваційна мережа штучного інтелекту для загального добробуту» (2018) • Канадський Форум штучного інтелекту провінції Квебек (2020) • Південнокорейський Форум з політики етики штучного інтелекту (2022)
Соціальний діалог і залучення недостатньо представлених груп населення	<ul style="list-style-type: none"> • Ініціатива «Відкритий діалог: штучний інтелект у Канаді» (2020) • Національний альянс штучного інтелекту Мексики (ANIA) (2023)
Сприяння використанню штучного інтелекту для забезпечення екологічної стійкості	<ul style="list-style-type: none"> • Німецька програма «Маяки штучного інтелекту для навколишнього середовища, клімату, природи та ресурсів» (2019) • Проєкт Європейської комісії «Destination Earth (DestinE)» для боротьби зі змінами клімату на основі використання моделі «цифрового двійника» Землі (2023)
Розширення можливостей та підвищення творчих здібностей людей	<ul style="list-style-type: none"> • Турецька програма «Виявлення раку молочної залози за допомогою штучного інтелекту» • Німецька виставка «AUTONOM – Виконавське мистецтво та штучний інтелект» (2021)

Джерело: складено на основі [11; 31; 32; 34; 35]

Так, програма французького уряду «IA Booster» спрямована на зменшення соціальної та економічної нерівності шляхом підтримки малих і середніх підприємств на основі проведення цифрової трансформації їх діяльності з використанням технологій штучного інтелекту [31]. Програма «Жінки у науці про дані та штучному інтелекті», започаткована у Великобританії Інститутом Алана Тюрінга, покликана сприяти зменшенню гендерної нерівності та розширенню можливостей участі жінок у сфері штучного інтелекту та науках про дані шляхом посилення співробітництва з політиками та зацікавленими сторонами [32; 33, с. 76]. Австралійський Національний центр штучного інтелекту

(NAIC) здійснює підвищення навичок з метою подолання регіональної, соціальної, економічної та гендерної нерівності [11].

Прикладом міждисциплінарного співробітництва є діяльність Канадського Форуму штучного інтелекту провінції Квебек, який використовує сучасні цифрові технології як інструмент прискорення економічного та соціального розвитку шляхом залучення широкого кола зацікавлених сторін для реалізації проєктів, спрямованих на підвищення міжнародної конкурентоспроможності постачальників технологій і систем штучного інтелекту та забезпечення їх етичного і соціально відповідального впровадження [34].

Німецька мережа «Громадянське кодування – інноваційна мережа штучного інтелекту для загального добробуту», яка була створена у 2018 р. міністерством праці та соціальної політики, міністерством у справах сім'ї, людей похилого віку, жінок і молоді та міністерством навколишнього середовища, покликана сприяти вирішенню широкого кола соціальних та екологічних завдань з використанням технологій штучного інтелекту на основі міжвідомчого співробітництва [35].

Національний альянс штучного інтелекту Мексики (ANIA) створено як відкритий інклюзивний простір для участі всіх зацікавлених сторін з метою аналізу можливостей і проблем, пов'язаних з використанням штучного інтелекту, зміцнення його екосистеми та оцінки впливу на забезпечення стійкого розвитку Мексики. Федеральне міністерство навколишнього середовища, охорони природи та ядерної безпеки Німеччини (BMUV) у 2019 р. започаткувало програму «Маяки штучного інтелекту для навколишнього середовища, клімату, природи та ресурсів», яка спрямована на підтримку використання технологій штучного інтелекту з метою розв'язання екологічних проблем. Більш широкі цілі у сфері використання штучного інтелекту з метою боротьби зі зміною клімату передбачено проєктом Європейської комісії «Destination Earth» (DestinE), відповідно до якого розроблено модель «цифрового двійника» Землі [11].

Набувають поширення урядові та приватні ініціативи щодо використання технологій штучного інтелекту з метою розширення можливостей людей у галузі медицини (Ізраїль, Португалія, США, Туреччина), освіти (Канада, Південна Корея, Чилі, Швеція), соціальної політики (Австрія, Великобританія, Німеччина Франція), культури (Німеччина) та інші.

2. Відповідно до принципу «Людиноцентричні цінності та справедливість» політичні ініціативи та рішення національних урядів повинні передбачати конкретні механізми та інструменти для реалізації цінностей, орієнтованих на людину, що дозволяють зменшити упередженість результатів і рекомендацій, забезпечують захист свобод,

прав, гідності та демократичних цінностей, а також конфіденційність даних, недискримінацію, різноманіття, соціальну справедливість і трудові права людей у разі використання відповідних технологій упродовж усього життєвого циклу систем штучного інтелекту. Конкретні приклади реалізації цих завдань урядами окремих країн світу наведено в табл. 2.

Таблиця 2

**Завдання політики та приклади реалізації вимог
принципу «Людиноцентричні цінності та справедливість»
в окремих країнах світу**

Завдання політики	Приклади реалізації
Захист прав і свобод людини, в т. ч. при використанні алгоритмічних систем для управління персоналом (HRIAs)	<ul style="list-style-type: none"> • Рекомендації Ради Європи про вплив алгоритмічних систем на права людини (CM/REC) (2020) • Рекомендації уряду Нідерландів «Оцінка впливу фундаментальних прав та алгоритмів» (FRAIA) (2021) • Проєкт Франції та Всесвітнього економічного форуму «Вирішення проблем з правами людини, пов'язаних з технологією розпізнавання обличчя» (2023)
Захист конфіденційності даних	<ul style="list-style-type: none"> • Мексиканські «Рекомендації щодо обробки персональних даних, отриманих у результаті використання штучного інтелекту» (2021) • Південнокорейський «Контрольний список захисту особистої інформації зі штучного інтелекту» (2022)
Знаки якості та сертифікати для просування цінностей, орієнтованих на людину	<ul style="list-style-type: none"> • «Знак якості штучного інтелекту» німецької асоціації штучного інтелекту (2019) • Мальтійська програма сертифікації штучного інтелекту (2019) • Турецький «Надійний знак якості штучного інтелекту» (2023)
Зменшення упередженості результатів і рішень, які надають системи штучного інтелекту	<ul style="list-style-type: none"> • «Огляд упередженості алгоритмічного прийняття рішень» Великобританії (2020) • Розслідування причин дискримінації в галузі штучного інтелекту Міністерства внутрішніх справ Нідерландів (2021) • Дослідницький семінар зі створення великих мовних моделей і наборів даних “BigScience Research Workshop” (2021–2022) • «Ініціатива штучного інтелекту та алгоритмічна справедливість» США (2023)

Джерело: складено на основі [11; 36; 37; 38; 40; 41]

Відповідно до Рекомендації CM/Rec(2020)1 Комітету міністрів Ради Європи, держави зобов'язані регулярно проводити оцінку впливу алгоритмічних систем на права людини, щоб виявити ризики

несприятливих результатів. Крім того, Комітет Ради Європи зі штучного інтелекту (CAI) працює над «Рамковою конвенцією зі штучного інтелекту, прав людини, демократії та верховенства закону» [36].

З метою управління ризиками для прав людини, які пов'язані з використанням алгоритмічних систем для управління персоналом (HRIAs), уряд Нідерландів розробив Рекомендації «Оцінка впливу фундаментальних прав та алгоритмів» (FRAIA) і вперше запровадив обов'язкове їх використання в органах державної влади країни. Одним із прикладів реалізації політичних рішень у сфері захисту прав і свобод людини є спільний проєкт Франції та Всесвітнього економічного форуму «Вирішення проблем з правами людини, пов'язаних з технологією розпізнавання обличчя», який ґрунтується на використанні принципів етичного та відповідального використання технологій штучного інтелекту та сприяє зниженню політичних і соціальних ризиків [11].

У Мексиці Національний інститут прозорості, доступу до інформації та захисту персональних даних (INAI) розробив «Рекомендації з обробки персональних даних, одержаних у результаті використання штучного інтелекту» [37]. Естонія, Туреччина, Великобританія та США розробляють нові технології підвищення конфіденційності. У Німеччині, Мальті та Туреччині запроваджуються знаки якості та сертифікати, які підтверджують, що рекомендовані до використання інструменти штучного інтелекту є етичними та орієнтовані на людину [11].

Найбільш активно питання алгоритмічної упередженості при використанні систем штучного інтелекту вивчаються у США. Так, Комісія США з рівних можливостей у сфері зайнятості розробила технічне керівництво з алгоритмічної справедливості для використання відповідних інструментів у процесі прийняття рішень щодо працевлаштування. Відповідно до цього керівництва, федеральні агентства, що використовують штучний інтелект в системах HRIAs, зобов'язані забезпечувати захист громадян від алгоритмічної дискримінації. Федеральна торгова комісія США (FTC) з метою уникнення незаконних або несправедливих дій і практик, пов'язаних із використанням алгоритмів і біометричних методів у комерційній діяльності, запровадила систему регулятивних заходів [38].

У 2020 р. Центр етики та інновацій даних Великобританії (CDEI) на основі аналізу впливу використання алгоритмічних інструментів на упередженість у процесі прийняття управлінських рішень при наданні фінансових послуг, наборі персоналу, а також у діяльності поліції та місцевого самоуправління розробив пропозиції щодо зниження цифрових ризиків [39, с. 58; 40].

Міжнародний відкритий дослідницький семінар з питань створення великих багатомовних моделей і наборів даних BigScience Research

Workshop, який було проведено у 2021–2022 рр. у Франції за участі понад 1000 дослідників із 60 країн світу, розробив рекомендації щодо етичного використання технологій штучного інтелекту та зниження рівня упередженості результатів нейронних мовних моделей [41].

3. Вимоги принципу «Прозорість і зрозумілість» ґрунтуються на тому, що прозорість і пояснюваність логіки та чітка інтерпретація процесів прийняття рішень є ключовими характеристиками відповідальних і безпечних систем штучного інтелекту, а тому передбачають необхідність реалізації широкого кола політичних ініціатив, заходів і практичних дій урядів щодо загального розуміння та забезпечення прозорості функціонування і відповідального розкриття інформації про проєктування та використання систем штучного інтелекту, надання інформації про фактори та процеси прийняття рішень з метою оскарження їх результатів і відшкодування збитків [42, с. 262]. Конкретні механізми та інструменти, які використовують окремі країни з метою забезпечення прозорості та зрозумілості функціонування систем штучного інтелекту, наведено в табл. 3.

Таблиця 3

Завдання політики та приклади реалізації вимог принципу «Прозорість і зрозумілість» в окремих країнах світу

Завдання політики	Приклади реалізації
Розкриття інформації про використання систем штучного інтелекту	<ul style="list-style-type: none"> • Закон Франції «Про цифрову республіку» (2016) • Директива ЄС «Про цифрові послуги» (DSA) (2022) • Англійський «Стандарт алгоритмічної прозорості запису» (2022) • Закон штату Нью-Йорк про електронний моніторинг (2022) • Англійське керівництво з юридичних зобов'язань роботодавців при спостереженні за працівниками (2022) • Японський Закон про прозорість цифрових платформ (DP Transparency Act) (2023)
Надання інформації про роботу штучного інтелекту	<ul style="list-style-type: none"> • Закон США «Про алгоритмічну підзвітність» (AAA) (2022) • Іспанський «Закон про пасажирів», який визначає вимоги роботодавців до алгоритмічної прозорості (2021) • Постанова уряду Китаю про алгоритмічні рекомендаційні системи (2022)
Надання інформації про фактори та процеси прийняття рішень і можливості відшкодування збитків від рішень, прийнятих на основі штучного інтелекту	<ul style="list-style-type: none"> • Загальний регламент ЄС із захисту даних (GDPR) (2018) • Законопроект Великобританії «Про захист даних та інформації» (2023) • Канадська Директива про автоматизоване прийняття рішень (2023)

Джерело: складено на основі [11; 23; 43]

У контексті реалізації етичних принципів використання штучного інтелекту прозорість означає, що на основі наданої інформації користувачі усвідомлюють факт своєї взаємодії з системою штучного інтелекту, а логіка прийняття рішень має чітку інтерпретацію та пояснюваність [11; 16, с. 285].

Відповідно до законопроекту ЄС «Про штучний інтелект», прозорість функціонування систем штучного інтелекту повинна забезпечуватися шляхом інформування користувачів про їхню взаємодію з такими системами. Для окремих випадків визначено особливі умови підтримки прозорості систем з високим рівнем ризику, а також таких, що використовуються для розпізнавання емоцій, біометричної категоризації, створення чи управління зображенням, аудіо, відео тощо. Крім того, Законом ЄС «Про цифрові послуги» (DSA) передбачено чіткі вимоги щодо підвищення рівня прозорості цифрових алгоритмів. Зокрема, великі онлайн-платформи (VLOP) та онлайн-пошукові системи (VLOSE) зобов'язані виявляти, здійснювати оцінку та сприяти зменшенню впливу системних ризиків, які можуть виникнути при проектуванні та функціонуванні їхніх сервісів і пов'язаних з ними алгоритмічних систем. Для виконання регуляторних і наглядових функцій створено Європейський центр алгоритмічної прозорості (ECAT), який здійснює оцінку відповідності алгоритмічних систем вимогам щодо управління ризиками, визначеним законом для VLOP і VLOSE [43].

Уряд Японії, згідно із законом «Про прозорість цифрових платформ» (DP Transparency Act), вимагає від постачальників цифрових платформ забезпечувати прозорість і справедливість транзакцій шляхом розкриття інформації про використання систем штучного інтелекту. Французький Закон «Про цифрову республіку» зобов'язує державні установи забезпечувати прозорість алгоритмічних інструментів, які використовує уряд, шляхом їх регулярної публікації, а Фінляндія та Нідерланди з метою оцінки використання алгоритмів у органах місцевої влади створили відкриті реєстри штучного інтелекту. Центр стандартів штучного інтелекту Великобританії розробив Стандарт алгоритмічної прозорості запису, який встановлює вимоги до розкриття інформації про використання алгоритмічних інструментів [43].

Значну увагу уряди окремих країн приділяють використанню штучного інтелекту на робочому місці. Так, у Великобританії, Іспанії, Італії, Канаді, Німеччині, США, Франції та інших країнах роботодавці, які використовують інструменти штучного інтелекту, зобов'язані інформувати своїх працівників про політику та методи електронного моніторингу їх діяльності.

З метою забезпечення прозорості використання систем штучного інтелекту канадський законопроект «Про штучний інтелект і дані»

зобов'язує компанії надавати чітку, достовірну та достатню інформацію про потенційні можливості алгоритмічних систем на загальнодоступному веб-сайті, а також зазначати потенційні ризики їх застосування. Законопроект США про алгоритмічну підзвітність (AAA) визначає вимоги до компаній, які використовують алгоритмічні системи для прийняття критично важливих рішень, що можуть мати значні юридичні або матеріальні наслідки для споживачів, зокрема, доступ до освіти, працевлаштування, медичного обслуговування, певних видів фінансових і комунальних послуг [23; 44, с. 46].

Загальний регламент ЄС щодо захисту даних (GDPR) передбачає можливість надання на запит громадян пояснень щодо прийнятих відносно них рекомендації або рішень, отриманих з використанням виключно автоматизованих систем, якщо ці рішення спричиняють негативні правові чи репутаційні наслідки. Схожі положення містяться у законодавстві Великобританії, Канади, Мексики, Норвегії та інших країн [43].

4. Відповідно до принципу «Надійність, захищеність і безпека» ініціативи національних урядів повинні сприяти забезпеченню належного управління ризиками, що виникають унаслідок використання систем штучного інтелекту, включаючи конфіденційність, цифрову безпеку та упередженість, а також розробляти та запроваджувати стандарти і правила ведення обліку даних.

Досвід окремих країн щодо забезпечення надійності, захищеності та безпеки функціонування технологій і систем штучного інтелекту представлено в табл. 4.

Прикладом ефективного забезпечення надійності та безпечності функціонування систем штучного інтелекту є методика алгоритмічної оцінки впливу (AIA), розроблена відповідно до канадського законодавства про автоматизоване прийняття рішень, використання якої допомагає виявити та оцінити потенційні соціальні, економічні, етичні та юридичні наслідки, пов'язані з проектом, системою, алгоритмом, рішенням, впливом, даними та заходами щодо пом'якшення ризиків. За рівнем ризику системи AIA класифікуються чотири рівні впливу: незначний (або відсутній), помірний, високий та дуже високий. Важливою умовою використання зазначеної методики є те, що оцінка впливу спочатку здійснюється до запуску системи, а потім регулярно проводиться впродовж усього її життєвого циклу [45]. Досвід Канади набув поширення в Уругваї та Мексиці, яка розробила «Посібник з принципів та аналізу впливу розробки і використання систем на основі штучного інтелекту в державному управлінні».

**Завдання політики та приклади реалізації вимог
принципу «Надійність, захищеність і безпека»
в окремих країнах світу**

Завдання політики	Приклади реалізації
Алгоритмічна оцінка впливу	<ul style="list-style-type: none"> • Канадська методика алгоритмічної оцінки впливу штучного інтелекту (AIA) (2023) • Мексиканський «Посібник з принципів та аналізу впливу розробки та використання систем на основі штучного інтелекту в державному управлінні» (2022) • Уругвайське «Дослідження алгоритмічного впливу (EIA)» (2022)
Ведення обліку даних для оцінки простежуваності	<ul style="list-style-type: none"> • «Національний словник даних» Туреччини (2019) • Президентський меморандум національної безпеки США (NSPM) «Захист переваг США в галузі штучного інтелекту та пов'язаних з ним критичних технологій» (2019)
Закони та нормативно-правові акти, що запобігають необґрунтованим ризикам для безпеки використання систем штучного інтелекту	<ul style="list-style-type: none"> • Законодавство Австрії, Великобританії, Данії, Китаю, Литви, Німеччини, США та Японії про автономне управління автомобілем

Джерело: складено на основі [11; 21; 43; 45]

Головними завданнями реалізації Меморандуму національної безпеки США «Захист переваг США в галузі штучного інтелекту та пов'язаних з ним критичних технологій» (NSPM) є захист критичних технологій та поліпшення доступу дослідників до федеральних даних, моделей і обчислювальних ресурсів [21]. Заслуговує на увагу досвід Туреччини щодо створення Національного словника даних, метою якого є формування національного реєстру даних і налагодження процесів управління та проведення моніторингу на основі національної архітектури інтеграції даних [11].

Упродовж останніх років у багатьох країнах набувають поширення законодавчі ініціативи щодо використання автономних транспортних засобів та автономного водіння. Розроблені закони та нормативні акти визначають умови використання автономних транспортних засобів і автоматизованих систем водіння, встановлюють вимоги до пересування дорогами загального користування тощо. Прикладами ефективних законодавчих ініціатив, що сприяють попередженню ризиків безпеки при використанні систем штучного інтелекту для автономних транспортних засобів і автономного водіння, є досвід Австрії, Великобританії, Данії, Китаю, Литви, Німеччини, США, Японії та інших країн [43].

5. Принцип «Підзвітність» передбачає, що організації та особи, які проєктують, впроваджують і використовують системи штучного інтелекту зобов'язані забезпечувати та нести відповідальність за правильне їх функціонування упродовж усього життєвого циклу, керуючись при прийнятті рішень відповідними законодавчими та нормативними документами [11].

Головними завданнями урядів і міжнародних організацій при забезпеченні належного рівня підзвітності використання систем штучного інтелекту є документування процесів упродовж усього терміну їх функціонування, розроблення кодексів етичної поведінки, стандартів та інші інструментів, а також створення та забезпечення ефективної діяльності наглядових органів.

Ініціативи урядів окремих країн передбачають розробку законодавчих і нормативних актів, кодексів етичної поведінки, створення незалежних наглядових органів, які перевіряють використання алгоритмів і умов функціонування систем, застосування інструментів документування та забезпечення підзвітності в процесі проєктування, впровадження і використання технологій і систем штучного інтелекту в бізнесі, державному управлінні, освіті, охороні здоров'я, науці, автономному водінні тощо (табл. 5).

Упродовж останніх років у сфері підтримки процесів документування рішень щодо проєктування, впровадження і використання технологій і систем штучного інтелекту, а також забезпечення їх підзвітності суттєво активізувалась діяльність міжнародних організацій, зокрема, Міжнародної організації зі стандартизації (ISO), Інституту інженерів з електротехніки та електроніки (IEEE), Міжнародного союзу електрозв'язку (ITU), Європейського інституту стандартів телекомунікацій (ETSI), Європейського комітету з електротехнічної стандартизації (CEN-CENELEC), Європейського альянсу штучного інтелекту та інших [11]. Їхні зусилля та ініціативи зосереджені, головним чином, на стандартизації технологій і систем штучного інтелекту, розробленні нових підходів до проєктування, оцінки впливу, відповідності та управління ризиками.

Найбільшу увагу питанням документування процесу функціонування систем штучного інтелекту упродовж усього їх життєвого циклу приділяють уряди США, ЄС та Канади. Так, відповідно до закону США «Про алгоритмічну звітність», з метою формування документації для автоматизованих систем прийняття рішень оператори систем штучного інтелекту зобов'язані підтримувати та оновлювати документацію щодо даних та іншої вхідної інформації, яка використовується для розробки, тестування та обслуговування алгоритмічних систем прийняття рішень [23]. Законопроект ЄС «Про штучний інтелект» зобов'язує агентів надавати національним регуляторам необхідну технічну документацію,

особливо щодо систем штучного інтелекту з високим рівнем ризику. Законопроект Канади «Про штучний інтелект і дані» зобов'язує компанії впроваджувати відповідні механізми управління, необхідні для документування процесів, а також дотримання юридичних зобов'язань щодо використання систем штучного інтелекту [20].

Таблиця 5

Завдання політики та приклади реалізації вимог принципу «Підзвітність» в окремих країнах світу

Завдання політики	Приклади реалізації
Розроблення законодавчих актів, які вимагають документального оформлення належного функціонування систем штучного інтелекту впродовж їх життєвого циклу	<ul style="list-style-type: none"> • Закон США «Про алгоритмічну звітність» (AAA) (2019) • Проект Закону Канади «Про штучний інтелект і дані» (AIDA) (2022) • Закон ЄС «Про штучний інтелект» (2024)
Розроблення кодексів етичної поведінки та запровадження практичних інструментів	<ul style="list-style-type: none"> • Кодекси етичної поведінки використання штучного інтелекту в державному управлінні та наданні соціальних послуг (Великобританія, Корея, Норвегія, Португалія, Франція) • Рекомендації щодо передової практики з інтеграції етики в розробку рішень штучного інтелекту в системі охорони здоров'я (Корея, Сінгапур, Франція) • Кодекси етичної поведінки з використання штучного інтелекту в автономному керуванні автомобілями (Австрія, Німеччина, Південна Корея) • Каталог інструментів і показників для штучного інтелекту ОЕСР (2023) • Портфель варіантів використання та онлайн сховище інструментів для безпечного використання штучного інтелекту (AI Assurance) Центру етики та інновацій даних Великобританії (CDEI) • Сінгапурська система перевірки штучного інтелекту (AI Verify) для тестування управління ризиками штучного інтелекту в компаніях
Створення незалежних органів нагляду для аудиту використання алгоритмів	<ul style="list-style-type: none"> • Органи нагляду за алгоритмами та функціонуванням моделей штучного інтелекту (Іспанія, Канада, Нідерланди, Норвегія)

Джерело: складено на основі [11; 19; 20; 23]

Значну увагу національні уряди приділяють створенню та забезпеченню ефективного функціонування в своїх країнах наглядових органів для перевірки та забезпечення прозорості використання

алгоритмів штучного інтелекту. Так, у 2022 р. в Іспанії було створено агентство з нагляду за використанням штучного інтелекту, а в 2024 р. – Орган алгоритмічного нагляду Норвегії. В Європейському Союзі створено Європейську раду зі штучного інтелекту. В Управлінні із захисту даних Нідерландів функціонує підрозділ, який здійснює нагляд за алгоритмами обробки персональних даних, проводить їх моніторинг, розробляє заходи для попередження дискримінації та упередженості, а також виявляє та проводить аналіз ризиків, на які вони можуть наражати користувачів [11].

Узагальнення світового досвіду свідчить про те, що найбільш нагальними питаннями для України у сфері управління штучним інтелектом на основі використання ціннісних принципів є прийняття закону «Про штучний інтелект» та оновлення відповідно до вимог Загального регламенту ЄС про захист даних (GDPR) Закону України «Про захист персональних даних», створення національного органу регулювання та нагляду за штучним інтелектом і даними, посилення державної підтримки процесів впровадження і використання штучного інтелекту, розроблення індикаторів для вимірювання його впливу на економічні та соціальні процеси, підвищення рівня обізнаності бізнесу та громадян щодо потенціалу, переваг і ризиків використання, а також вдосконалення управління ризиками, на які можуть наражати рішення та рекомендації, надані системами штучного інтелекту.

Реалізація цих заходів сприятиме посиленню цифрового суверенітету країни, прискоренню темпів економічного зростання та повного відновлення, а також суттєвому підвищенню рівня міжнародної конкурентоспроможності.

Висновки

Необхідність підвищення ефективності використання штучного інтелекту в різних галузях економіки та сферах діяльності людей, забезпечення цифрового суверенітету країн та зниження рівня ризиків, на які системи штучного інтелекту можуть наражати органи влади, бізнес і громадян, актуалізує необхідність розроблення дієвих методів та інструментів регулювання та управління на основі використання принципів, орієнтованих на дотримання законних прав і свобод людей та демократичних цінностей. Головною метою такого регулювання повинно бути забезпечення етичного, безпечного, надійного та соціально відповідального проектування, впровадження і використання технологій і систем штучного інтелекту.

Міжнародні організації та більшість країн світу сьогодні активно розробляють керівні принципи, стандарти, нормативно-правову базу та правила безпечного і соціально відповідального впровадження і

використання технологій і систем штучного інтелекту шляхом формування відповідних політик і стратегій та ґрунтуючись на галузевих або вертикальних підходах до управління, щоб зробити функціонування систем штучного інтелекту прозорим, безпечним і підзвітним.

До головних ціннісних принципів розроблення та використання технологій і систем штучного інтелекту належать їх орієнтація на людиноцентричні цінності, забезпечення інклюзивного зростання, стійкого розвитку та добробуту, прозорості та безпечності. Практична реалізація цих принципів передбачає розробку відповідних заходів державної політики щодо захисту прав і свобод людини, зменшення економічної та соціальної нерівності, розширення соціального діалогу, підтримки екологічної стійкості, розширення можливостей людей, в тому числі й на ринку праці, зниження економічних, соціальних і політичних ризиків, а також зменшення упередженості результатів і рішень, які надають системи штучного інтелекту.

Інструментами реалізації завдань державної політики щодо управління штучним інтелектом на основі використання ціннісних принципів є розроблення відповідних законодавчих і нормативних актів, створення незалежних органів нагляду, захист конфіденційності даних, розкриття інформації щодо прийняття рішень, організація належного документування процесів функціонування систем штучного інтелекту упродовж їх життєвого циклу та інші.

Завдання України у сфері управління штучним інтелектом повинні передбачати розроблення відповідних нормативно-правових актів, створення органів державного регулювання та контролю, підвищення обізнаності бізнесу і громадян щодо потенціалу та управління ризиками, на які їх можуть наражати рішення, надані системами штучного інтелекту. Подальші дослідження передбачають забезпечення міжнародної сумісності систем штучного інтелекту та розроблення індикаторів оцінки їх впливу на економічні та соціальні процеси.

Список використаних джерел:

1. AI regulation: a pro-innovation approach. UK Government. 2023. URL: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>
2. Recommendation of the Council on Artificial Intelligence. OECD. 2019. URL: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>
3. What is artificial intelligence and how is it used? European Parliament. 2023. URL: <https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used>
4. Концепція розвитку штучного інтелекту і в Україні. Розпорядження Кабінету Міністрів України від 02.12.2020 р. № 1556-р. URL: <https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%80#Text>

5. Дорожня карта з регулювання штучного інтелекту в Україні. Міністерство цифрової трансформації України. 2023. URL: <https://thedigital.gov.ua/storage/uploads/files/page/community/docs/>
6. Recommendation on the Ethics of Artificial Intelligence. UNESCO. SHS/BIO/PI/2021/1. 2021. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
7. Hiroshima Leaders' Communiqué. Ministry of Foreign Affairs of Japan. G7. 2023. URL: https://www.mofa.go.jp/ms/g7hs_s/page1e_000690.html
8. Hiroshima AI Process G7. Digital & Tech Ministers' Statement. 2023. URL: https://www.soumu.go.jp/main_content/000900470.pdf
9. National Artificial Intelligence Initiative Act of 2020. H.R. 6216. 2020. URL: https://democrats-science.house.gov/imo/media/doc/AI_initiative_SST.pdf
10. AI Strategy 2022. Japanese Government. 2022. URL: <https://www8.cao.go.jp/cstp/ai/aistrategy2022en.pdf>
11. The state of implementation of the OECD AI Principles four years on. OECD Artificial Intelligence Papers. 2023. No. 3. DOI: <https://doi.org/10.1787/835641c9-en>
12. Міщенко В.І. Стратегічне управління процесами цифрової трансформації економіки. *Економіка України*. 2022. № 1. С. 67–81.
13. Науменкова С.В., Міщенко В.І., Міщенко С.В. Цифрова фінансова інклюзія для досягнення Цілей сталого розвитку. *Вісник Київського національного університету імені Тараса Шевченка. Економіка*. 2023. № 1. С. 102–112. DOI: <https://doi.org/10.17721/1728-2667.2023/222-1/13>
14. Про Національну комісію з питань захисту персональних даних та доступу до публічної інформації. Проект Закону України від 06.02.2024 р. № 3562-IX. URL: https://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=72992
15. Гриценко А. А. та ін. Відновлення та реконструкція повоєнної економіки України: наукова доповідь. НАН України, ДУ «Ін-т екон. та прогноз. НАН України». Київ : 2022. 305 с. URL: <http://ief.org.ua/wp-content/uploads/2022/12/Vidnovlennja-ta-rekonstrukcija-povojennoj-ekonomiki.pdf>
16. Міщенко В. І. Механізми регулювання процесів цифровізації для забезпечення національно укорінені стійкості економічного розвитку. *Економічний простір*. 2024. № 189. С. 283–290. DOI: <https://doi.org/10.32782/2224-6282/189-50>
17. Науменкова С., Міщенко С., Тіщенко Є. Фінансові інструменти забезпечення від політичних ризиків у проектному фінансуванні. *Вісник Київського національного університету імені Т. Шевченка*. 2016. № 10. С. 6–17.
18. La stratégie nationale de recherche en intelligence artificielle. Cour des Comptes. 2023. URL: <https://www.ccomptes.fr/fr/publications/la-strategie-nationale-de-recherche-en-intelligence-artificielle>
19. The Artificial Intelligence and Data Act (AIDA). Government of Canada. 2022. URL: <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document>
20. Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI. European Parliament. Press Releases. 09-12-2023. URL: <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>
21. Maintaining American Leadership in Artificial Intelligence. US Presidential Office. 2019. URL: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>
22. Blueprint for an AI Bill of Rights. US Government. 2022. URL: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>

23. Algorithmic Accountability Act (AAA). US Congress. 2022. URL: <https://www.congress.gov/bill/117th-congress/house-bill/6580>
24. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. January 2023. URL: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
25. Principles for Policy, Regulation and Ethics in AI: technical paper (annex to policy paper). Ministry of Innovation, Science and Technology and Ministry of Justice, the Office of legal counsel and legislative affairs. 2022. URL: <https://www.gov.il/BlobFolder/rfp/061122/he/professional-letter.pdf>
26. Social Principles of Human-Centric AI. Japanese Government. 2019. URL: <https://www8.cao.go.jp/cstp/stmain/aisocialprinciples.pdf>
27. Opinion on Strengthening the Ethics and Governance in Science and Technology. Chinese Government. 2022. URL: http://www.gov.cn/zhengce/2022-03/20/content_5680105.htm
28. Міщенко В.І. Механізми регулювання обміну даними. *Причорноморські економічні студії*. 2022. № 75. С. 37–45. DOI: <https://doi.org/10.32843/bses.75-6>
29. Дзюблюк О. В. та ін. Теорія і практика грошового обігу та банківської справи в умовах глобальної фінансової нестабільності: монографія. Тернопіль : ФОП Осадца Ю.В., 2017. 298 с.
30. Міщенко С.В., Науменкова С.В., Міщенко В.І. Управління ризиками зеленого фінансування в Україні. *Вісник Одеського національного університету ім. І.І. Мечникова. Економіка*. 2023. Т. 28. № 3 (97). С. 61–69. DOI: <https://doi.org/10.32782/2304-0920/3-97-11>
31. OECD Business and Finance Outlook 2021: AI in Business and Finance. OECD. 2021. URL: <https://www.oecd.org/finance/oecd-business-and-finance-outlook-26172577.htm>
32. Women in Data Science and AI. The Alan Turing Institute. 2023. URL: <https://www.turing.ac.uk/research/research-projects/women-data-science-and-ai-new>
33. Міщенко В.І., Міщенко С. В. Маркетинг цифрових інновацій на ринку банківських послуг. *Фінансовий простір*. 2018. № 1 (29). С. 75–79.
34. Artificial intelligence will change the world. Forum IA Quebec. 2023. URL: <https://forumia.quebec/en/>
35. Über uns. Civic Coding. 2018. URL: <https://www.civic-coding.de/ueber-uns>
36. Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. Council of Europe. 2020. URL: https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154
37. Recomendaciones para el tratamiento de datos personales derivado del uso de la Inteligencia Artificial. INAI. 2022. URL: <https://home.inai.org.mx/wpcontent/documentos/DocumentosSectorPublico/RecomendacionesPDP-IA.pdf>
38. Database of national AI policies & strategies. OECD.AI 2023. URL: <https://oecd.ai/en/dashboards/overview>
39. Міщенко В. І. Управління кібербезпекою в системі забезпечення національно укоріненої стійкості економічного розвитку. *Економічна теорія*. 2023. № 1. С. 47–72. DOI: <https://doi.org/10.15407/etet2023.01.047>
40. Review into bias in algorithmic decision-making. UK. 2020. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957259/Review_into_bias_in_algorithmic_decision-making.pdf
41. A one-year long research workshop on large multilingual models and datasets. BigScience. 2024. URL: <https://bigscience.huggingface.co/>

42. Mishchenko V., Naumenkova S., Ivanov V., Tishchenko I. Special aspects of using hybrid financial tools for project risk management in Ukraine. *Investment Management and Financial Innovations*. 2018. № 15 (2). P. 257–266. DOI: [http://dx.doi.org/10.21511/imfi.15\(2\).2018.23](http://dx.doi.org/10.21511/imfi.15(2).2018.23)
43. Live data: investments in AI and data. OECD.AI. 2023. URL: <https://oecd.ai/en/data?selectedArea=investments-in-ai-and-data>
44. Mishchenko S., Naumenkova S., Mishchenko V., Ivanov V., Lysenko R. Growing discoordination between monetary and fiscal policies in Ukraine. *Banks and Bank Systems*. 2019. № 14 (2). P. 40–49. DOI: [https://doi.org/10.21511/bbs.14\(2\).2019.04](https://doi.org/10.21511/bbs.14(2).2019.04)
45. Public Awareness Working Group. Government of Canada. 2023. URL: <https://isedisde.canada.ca/site/advisory-council-artificial-intelligence/en/public-awareness-workinggroup>

References:

1. UK Government. (2023) AI regulation: a pro-innovation approach. Available at: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>
2. OECD. (2019) Recommendation of the Council on Artificial Intelligence. Available at: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>
3. EP. (2023) What is artificial intelligence and how is it used? Available at: <https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used>
4. KMU (2020) Kontseptsija rozvytku shtuchnoho intelektu v Ukraini [The Concept of Artificial Intelligence Development in Ukraine]. No. 1556-p. Available at: <https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%80#Text> (in Ukrainian)
5. MTsTU (2023) Dorozhna karta z rehulyuvannya shtuchnoho intelektu v Ukraini [Roadmap for the regulation of artificial intelligence in Ukraine]. Available at: <https://thedigital.gov.ua/storage/uploads/files/page/community/docs/> (in Ukrainian)
6. UNESCO. (2021) Recommendation on the Ethics of Artificial Intelligence. SHS/BIO/PI/2021/1. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
7. Ministry of Foreign Affairs of Japan. G7. (2023) Hiroshima Leaders' Communiqué. Available at: https://www.mofa.go.jp/ms/g7hs_s/page1e_000690.html
8. G7. (2023) Hiroshima AI Process G7. Digital & Tech Ministers' Statement. 2023. Available at: https://www.soumu.go.jp/main_content/000900470.pdf
9. NAII (2020) National Artificial Intelligence Initiative Act of 2020. H.R. 6216. Available at: https://democrats-science.house.gov/imo/media/doc/AI_initiative_SST.pdf
10. Japanese Government. (2022) AI Strategy 2022. Available at: <https://www8.cao.go.jp/cstp/ai/aistratagy2022en.pdf>
11. OECD. (2023) The state of implementation of the OECD AI Principles four years on. Artificial Intelligence Papers. No. 3. Available at: <https://doi.org/10.1787/835641c9-en>
12. Mishchenko V.I. (2022) Stratehichne upravlinnja protsesamy tsyfrovoy transformatsii ekonomiky [Strategic management of the processes of digital transformation of the economy]. *Ekonomika Ukrainy*, no. 1, pp. 67–81 (in Ukrainian)
13. Naumenkova S.V., Mishchenko V.I., Mishchenko S.V. (2023) Tsyfrova finansova inkluzija dlja dosjahnennja Tsiley staloho rozvytku [Digital Financial Inclusion to Achieve the Sustainable Development Goals]. *Visnyk Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. Ekonomika*, no. 1, pp. 102–112. DOI: <https://doi.org/10.17721/1728-2667.2023/222-1/13> (in Ukrainian)

14. VRU. (2024) Pro Natsionalnu komisiju z pytan zakhystu personalnykh danykh ta dostupu do publichnoi informstsii [About the National Commission on Personal Data Protection and Access to Public Information]. No. 3562-IX. Available at: https://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=72992 (in Ukrainian)
15. Grytsenko A.A. et al. (2022) Vidnovlennia ta rekonstruktsiia povoiennoi ekonomiky Ukrainy: naukova dopovid [Restoration and reconstruction of the post-war economy of Ukraine: scientific report]. Kyiv: In-t ekon. ta prohnouzuv. NAN Ukrainy. Available at: <http://ief.org.ua/wp-content/uploads/2022/12/Vidnovlennja-ta-rekonstruktsja-povojennoji-econo-miky.pdf> (in Ukrainian)
16. Mishchenko V.I. (2024) Mekhanizmy rehulyvannja protsesiv tsyfrovizatsii dlja zabezpechennja natsionalno ukorinenoi stiykosti ekonomichnoho rozvytku [Mechanisms for regulating digitalization processes to ensure nationally rooted sustainability of economic development]. *Ekonomichniy prostir*, no. 189, pp. 283–290. DOI: <https://doi.org/10.32782/2224-6282/189-50> (in Ukrainian)
17. Naumenkova S., Mishchenko S., Tishchenko I. (2016) Finansovi instrumenty ubezpechennja vid politychnykh ryzykiv u proektnomu finansuvanni [Financial Instruments for Safeguarding Against Political Risks in Project Financing]. *Visnyk Kyivskoho natsionalnoho universytetu imeni Tarasa Shvchenka. Ekonomika*, no. 10, pp. 6–17 (in Ukrainian)
18. Cour des Comptes. (2023) La stratégie nationale de recherche en intelligence artificielle. Available at: <https://www.ccomptes.fr/fr/publications/la-strategie-nationale-de-recherche-en-intelligence-artificielle>
19. Government of Canada. (2022) The Artificial Intelligence and Data Act (AIDA). Available at: <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document>
20. European Parliament. (2023) Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI. Press Releases. Available at: <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>
21. US Presidential Office. (2019) Maintaining American Leadership in Artificial Intelligence. Available at: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>
22. US Government. (2022) Blueprint for an AI Bill of Rights. Available at: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>
23. US Congress. (2022) Algorithmic Accountability Act (AAA). Available at: <https://www.congress.gov/bill/117th-congress/house-bill/6580>
24. NIST (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0). AI 100-1. Available at: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
25. Ministry of Innovation, Science and Technology and Ministry of Justice, the Office of legal counsel and legislative affairs. (2022) Principles for Policy, Regulation and Ethics in AI: technical paper (annex to policy paper). Available at: <https://www.gov.il/BlobFolder/rfp/061122/he/professional-letter.pdf>
26. Japanese Government. (2019) Social Principles of Human-Centric AI. Available at: <https://www8.cao.go.jp/cstp/stmain/aisocialprinciples.pdf>
27. Chinese Government. (2022) Opinion on Strengthening the Ethics and Governance in Science and Technology. Available at: http://www.gov.cn/zhengce/2022-03/20/content_5680105.htm
28. Mishchenko V.I. (2022) Mekhanizmy rehulyvannja obminu danymy [Mechanisms for regulating data exchange]. *Prychornomorski ekonomichni studii*, no. 75, pp. 37–45. DOI: <https://doi.org/10.32843/bses.75-6> (in Ukrainian)

29. Dzyublyuk O.V. et al. (2017) Teoriia i praktyka hroshovoho obihu ta bankivskoi spravy v umovakh globalnoi nestabilnosti [Theory and practice of monetary circulation and banking in the conditions of global financial instability]. Ternopil : Osadtsa Yu.V. (in Ukrainian)
30. Mishchenko S.V., Naumenkova S.V., Mishchenko V.I. (2023) Upravlinnja ryzykamy zelenoho finansuvannja v Ukraini [Green Finance Risk Management in Ukraine]. *Visnyk Odeskoho natsionalnogo universytetu im. I.I. Mechnykova*, no. 28, 3 (97), pp. 61–69. DOI: <https://doi.org/10.32782/2304-0920/3-97-11> (in Ukrainian)
31. OECD. (2021) OECD Business and Finance Outlook 2021: AI in Business and Finance. Available at: <https://www.oecd.org/finance/oecd-business-and-finance-outlook-26172577.htm>
32. The Alan Turing Institute. (2023) Women in Data Science and AI. Available at: <https://www.turing.ac.uk/research/research-projects/women-data-science-and-ai-new>
33. Mishchenko V.I., Mishchenko S.V. (2018) Marketynh tsyfrovyykh innovatsiy na rynku bankivskukh posluh [Marketing of Digital Innovations in the Banking Services Market]. *Finansovyj prostir*, no. 1 (29), pp. 75–79 (in Ukrainian)
34. Forum IA Quebec. (2023) Artificial intelligence will change the world. Available at: <https://forumia.quebec/en/>
35. Civic Coding. (2018) Über uns. Available at: <https://www.civic-coding.de/ueber-uns>
36. Council of Europe. (2020) Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. Available at: https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154
37. INAI. (2022) Recomendaciones para el tratamiento de datos personales derivado del uso de la Inteligencia Artificial. Available at: <https://home.inai.org.mx/wpcontent/documentos/DocumentosSectorPublico/RecomendacionesPDP-IA.pdf>
38. OECD.AI. (2023) Database of national AI policies & strategies. Available at: <https://oecd.ai/en/dashboards/overview>
39. Mishchenko V.I. (2023) Upravlinnja kiberbezpekoyu v systemi zabezpechennja natsionalno ukorinenoi stiykosti ekonomichnogo rozvytku [Cybersecurity Management in the System of Ensuring Nationally Rooted Sustainability of Economic Development]. *Ekonomichna teorija*, no. 1, pp. 47–72. DOI: <https://doi.org/10.15407/etet2023.01.047> (in Ukrainian)
40. UK. (2020) Review into bias in algorithmic decision-making. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957259/Review_into_bias_in_algorithmic_decision-making.pdf
41. BigScience. (2024) A one-year long research workshop on large multilingual models and datasets. Available at: <https://bigscience.huggingface.co/>
42. Mishchenko V., Naumenkova S., Ivanov V., Tishchenko I. (2018) Special aspects of using hybrid financial tools for project risk management in Ukraine. *Investment Management and Financial Innovations*, no. 15 (2), pp. 257–266. DOI: [http://dx.doi.org/10.21511/imfi.15\(2\).2018.23](http://dx.doi.org/10.21511/imfi.15(2).2018.23)
43. OECD.AI. (2023) Live data: investments in AI and data. Available at: <https://oecd.ai/en/data?selectedArea=investments-in-ai-and-data>
44. Mishchenko S., Naumenkova S., Mishchenko V., Ivanov V., Lysenko R. (2019) Growing discoordination between monetary and fiscal policies in Ukraine. *Banks and Bank Systems*, no. 14 (2), pp. 40–49. DOI: [https://doi.org/10.21511/bbs.14\(2\).2019.04](https://doi.org/10.21511/bbs.14(2).2019.04)
45. Government of Canada. (2023) Public Awareness Working Group. Available at: <https://isedisde.canada.ca/site/advisory-council-artificial-intelligence/en/public-awareness-workinggroup>