

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ДЕРЖАВИ І ПРАВА ІМЕНІ В.М. КОРЕЦЬКОГО

Кваліфікаційна наукова праця на правах рукопису

УДК 341.1

ІГНАТУША ВІТАЛІЙ ВІКТОРОВИЧ

ДИСЕРТАЦІЯ

**СТАНОВЛЕННЯ ТА РОЗВИТОК МІЖНАРОДНО-ПРАВОВОГО
РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ**

Спеціальність 293 «Міжнародне право»
Галузь знань 29 «Міжнародні відносини»

Подається на здобуття ступеня доктора філософії (PhD)

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів авторів мають посилання на відповідне джерело.

_____ Ігнатуша В. В.

Науковий керівник: Проценко Ірина Миколаївна, кандидат юридичних наук.

Київ - 2026

АНОТАЦІЯ

Ігнатуша В. В. Становлення та розвиток міжнародно-правового регулювання штучного інтелекту. – *Кваліфікаційна наукова праця на правах рукопису.*

Дисертація на здобуття ступеня доктора філософії за спеціальністю 293 «Міжнародне право» (29 – Міжнародні відносини). – Інститут держави і права імені В. М. Корецького Національної академії наук України, Київ, 2026.

Дисертація є комплексним дослідженням теоретико-прикладних засад міжнародно-правового регулювання штучного інтелекту (далі – ШІ) на універсальному, міжрегіональному та регіональному рівнях. У роботі висвітлено генезу та еволюцію міжнародно-правового регулювання ШІ, розкрито поняття ШІ, досліджено його правову природу, основні властивості та здійснено загальний аналіз специфіки його міжнародно-правового регулювання. Особливу увагу приділено питанню правового статусу ШІ: обґрунтовано, що чинне міжнародне право не наділяє ШІ самостійним правовим статусом, а кваліфікує його як об'єкт міжнародно-правового регулювання, а не як суб'єкт. Проаналізовано ключові міжнародно-правові акти та акти права ЄС, що визначають правове положення ШІ та регулюють його розроблення, розгортання та використання. Особливу увагу зосереджено на тому, що сучасне міжнародно-правове регулювання ШІ ґрунтується не лише на формально-юридичних приписах, а й на етичних та ціннісних засадах – передусім на людиноцентричному підході, повазі до людської гідності, прав людини та основоположних свобод. Також проаналізовано стан, тенденції і закономірності формування багаторівневої системи міжнародно-правового регулювання цієї технології, зокрема: конвергенцію принципів регулювання навколо спільного ціннісного ядра (людиноцентричності, прозорості, підзвітності, недискримінації) за одночасної фрагментації регуляторних підходів, екстериторіальний вплив права ЄС через «Брюссельський ефект» та поширення ризикоорієнтованої моделі регулювання.

Актуальність дисертаційної роботи зумовлена стрімким розвитком технологій ШІ та безпрецедентним прискоренням нормотворчої активності у цій сфері, зокрема ухваленням у 2024 році першого міжнародного договору щодо ШІ – Рамкової конвенції Ради Європи про штучний інтелект та права людини, демократію та верховенство права (CETS № 225) – і першого горизонтального акта ЄС щодо ШІ, – Регламенту (ЄС) 2024/1689 (Акту про ШІ). Незважаючи на зростання кількості нормативних актів, спрямованих на регулювання ШІ, доктринальне осмислення правового статусу ШІ та його регулювання залишається фрагментарним, а комплексне дослідження, яке б охоплювало всю багаторівневу систему міжнародно-правового регулювання ШІ – від універсального до регіонального рівня – та його еволюцію від «м'якого» до «твердого» права, у вітчизняній науці міжнародного права відсутнє.

У першому розділі дисертаційного дослідження висвітлено історичну траєкторію розвитку концепції штучного інтелекту від Дартмутської конференції 1956 року, для проведення якої Дж. Маккарті запропонував сам термін «штучний інтелект», до сучасного етапу інтенсивної нормотворчої діяльності. Встановлено, що міжнародно-правове регулювання ШІ формувалося нелінійно: ШІ пройшов шлях від суто технічного феномену через етичні принципи та акти «м'якого права» до юридично обов'язкових норм. Розкрито проблему відсутності єдиного визначення штучного інтелекту в міжнародному праві та його науці, що породжує невизначеність меж предмета регулювання та ускладнює узгодження зобов'язань держав. На підставі аналізу міжнародно-правових актів та наукових джерел запропоновано авторську класифікацію концептуальних підходів до визначення ШІ: технологічного, антропоцентричного, функціонального, системного, еволюційного, цілеорієнтованого та інтегративного. Сформульовано авторську дефініцію штучного інтелекту як комплексної соціотехнічної системи, що характеризується функціональною здатністю до інференції (виведення висновків з даних), операційною автономністю різного ступеня, адаптивністю, генеративною спроможністю, транскордонним характером застосування, потенціалом подвійного

використання та створенням юридично значущих наслідків для сфери прав людини, міжнародного правопорядку та безпеки. Ця дефініція втілює інтегративний підхід, забезпечуючи технологічну нейтральність та придатність для використання у міжнародно-правових актах різного рівня.

Окреме місце у розділі посідає питання правового статусу ШІ. Обґрунтовано, що чинне міжнародне право не наділяє ШІ самостійним правовим статусом – на відміну від держави та міжнародної організації як загальновизнаних суб'єктів міжнародного права. Натомість обґрунтовано правову природу ШІ як об'єкта міжнародно-правового регулювання *sui generis*, якому притаманні поєднання матеріальних і нематеріальних складових, об'єктних і суб'єктних рис, динамічність та еволюційність, а також транскордонний характер застосування, потенціал подвійного використання, автономність різного ступеня, здатність до самовдосконалення та суттєвий вплив на міжнародні відносини та їх правове регулювання. Доведено, що фундаментальні відмінності ШІ від традиційних об'єктів міжнародно-правового регулювання зумовлюють необхідність формування спеціального правового режиму, який поєднує елементи різних галузей міжнародного права.

Встановлено, що міжнародно-правове регулювання ШІ характеризується фрагментацією регуляторних підходів на всіх рівнях та водночас стійкою конвергенцією принципів навколо спільного ціннісного ядра: людиноцентричного підходу, прозорості та пояснюваності, підзвітності та людського нагляду, безпеки та надійності, інклюзивності та недискримінації, захисту приватності та даних. Підкреслено, що саме ці етичні та ціннісні орієнтири, а не лише формальні юридичні приписи, формують ціннісно-нормативний фундамент сучасного регулювання ШІ. Розкрито специфіку співвідношення «м'якого» та «твердого» права у сфері міжнародно-правового регулювання ШІ, зокрема виокремлено чотири функції «м'якого права»: кристалізація консенсусу, правовий експеримент, легітимізація та підготовка підґрунтя для юридично обов'язкового регулювання. Водночас зроблено

висновок про недостатність самого лише «м'якого права» для ефективного регулювання технологій з масштабними та потенційно необоротними наслідками.

У другому розділі дисертаційного дослідження на підставі системного аналізу 23 резолюцій Генеральної Асамблеї ООН за період 2015–2025 років систематизовано дванадцять принципів, що формують нормативну основу міжнародно-правового регулювання ІІІ на універсальному рівні. Запропоновано оригінальну хронологію розвитку нормотворчої активності ГА ООН у сфері ІІІ та виявлено наступні закономірності: утвердження підходу на основі «життєвого циклу» систем ІІІ як базової регуляторної рамки, стандартизацію тріади «безпечні, захищені та надійні системи ІІІ» та позиціонування Глобального цифрового договору як точки конвергенції правового регулювання ІІІ на рівні ООН.

Також у другому розділі доведено, що діяльність спеціалізованих установ ООН – ЮНЕСКО, МСЕ, ВООЗ, ВОІВ та МОП – формує багатовимірну систему регулювання ІІІ з «градацією нормативності» у вигляді «концентричних кіл»: від актів з найвищим ступенем політичної легітимності, але водночас із найвищою абстрактністю, до галузевих стандартів, що стають де-факто обов'язковими завдяки вимогам інтероперабельності. Ця система являє собою поліцентричний інституційний ландшафт без чіткої ієрархії, в якому акти різного рівня співіснують та взаємно підсилюють одне одного.

Крім зазначеного, у другому розділі на підставі авторських критеріїв обґрунтовано доцільність виокремлення міжрегіонального рівня як самостійного предмета наукового аналізу. Встановлено, що ОЕСР виступає центральною ланкою міжрегіонального співробітництва у сфері регулювання ІІІ, з результатами роботи якої корелюють свою діяльність такі міждержавні формати, як G7, G20 та GPAI. Запропоновано та обґрунтовано концепцію «ітеративного нормотворення» як моделі формування актів «м'якого права» у сфері міжнародно-правового регулювання ІІІ, за якої акт функціонує як «живий документ», що проходить повторювані цикли ухвалення, апробації, оцінки ефективності та перегляду у відповідь на технологічні

зміни. Практичним втіленням цієї моделі є оновлення Рекомендації Ради ОЕСР щодо штучного інтелекту та систематичний перегляд Настанови АСЕАН з управління ШІ.

На основі проведеного дослідження виявлено та систематизовано ключові тенденції розвитку міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях: прискорення нормотворчої діяльності, перехід від фрагментарного до системного регулювання, інституціоналізація регулювання, розширення предмета та поглиблення деталізації регулювання, а також формування моделі «ітеративного нормотворення». Водночас виявлено сфери, які залишаються поза межами системного регулювання: запобігання негативному впливу ШІ на довкілля; відповідальність суб'єктів приватного права за шкоду, заподіяну функціонуванням систем ШІ; виклики, пов'язані з генеративним ШІ; вплив ШІ на формування світогляду людини та її когнітивну автономію; проблематика загального штучного інтелекту й довгострокових екзистенційних ризиків; а також воєнне використання ШІ.

У третьому розділі дисертаційного дослідження встановлено, що саме на регіональному рівні розпочався перехід від декларування принципів міжнародно-правового регулювання ШІ до формулювання юридично обов'язкових норм, наразі реалізований лише в рамках Європейського Союзу через Регламент (ЄС) 2024/1689 (Акт про ШІ) та під егідою Радою Європи через Рамкову конвенцію про штучний інтелект та права людини, демократію та верховенство права (CETS № 225). Запропоновано періодизацію формування правової бази ЄС у сфері регулювання ШІ та виявлено її системні закономірності: дуалізм ринкового та ціннісного обґрунтування регулювання, поступове послаблення механізмів забезпечення відповідності та реактивний характер регулювання. Обґрунтовано, що домінування ризикоорієнтованого підходу в сучасному міжнародно-правовому регулюванні ШІ відображає застосування принципу пропорційності та являє собою синтез запобіжного підходу (принципу обережності) з прагненням до стимулювання інновацій, що дозволяє забезпечувати баланс між захистом прав людини та інтересами технологічного розвитку. Найповнішим втіленням цього підходу є Акт

про ШІ, що запроваджує чотирирівневу класифікацію систем ШІ за ступенем ризику (неприйнятний, високий, обмежений та мінімальний) і диференціює обсяг регуляторних вимог відповідно до рівня потенційної шкоди для прав людини, безпеки та основоположних цінностей.

На підставі порівняльного аналізу нульового проєкту (англ. - zero draft) та фінального тексту названої Рамкової конвенції Ради Європи встановлено тенденцію до систематичного звуження обсягу зобов'язань держав, що відбулося в ході міжурядових переговорів. Це дає підстави для висновку про існування такої загальної закономірності: на етапі переговорів правозахисний потенціал юридично обов'язкових актів щодо регулювання ШІ послідовно зменшується під впливом як держав, так і суб'єктів приватного права. Водночас доведено взаємодоповнюваність двох європейських актів різної правової природи – Рамкової конвенції, що реалізує правозахисний підхід, та Акту про ШІ, який реалізує ризикоорієнтований підхід. Досліджено також адаптивний потенціал Конвенції про захист прав людини і основоположних свобод 1950 року (Європейської конвенції з прав людини) до проблематики ШІ, в результаті чого встановлено, що практика Європейського суду з прав людини поступово формує прецедентну базу, яка адаптується до цифрових реалій, однак ця практика має переважно реактивний характер і не може замінити проактивне правове регулювання ШІ, здійснюване через Рамкову конвенцію та Акт про ШІ.

Порівняльний аналіз підходів до правового регулювання ШІ, вироблених регіональними організаціями та об'єднаннями держав інших регіонів – АСЕАН, АТЕС, ШОС, Африканським Союзом, ОАД, Радою співробітництва арабських держав Перської затоки та АЛЕКСО, – дозволив ідентифікувати наступні ключові закономірності: «спектр юридичної обов'язковості», за якого «тверде право» залишається феноменом, притаманним виключно європейському регіону; «аксіологічний консенсус при регуляторному розмаїтті», за якого кожен регіон привносить у глобальний дискурс власні цінності та регуляторні підходи – від філософії Ubuntu в актах Африканського Союзу та ісламських етичних концепцій в

ініціативах держав Перської затоки до консенсусної моделі АСЕАН; диференційований характер «Брюссельського ефекту»; «парадокс глобального регулювання ШІ», за якого регіони з найбільшими потребами у захисті прав людини мають найменші інституційні спроможності; формування «регуляторного полілогу» між різними регіонами.

У результаті проведеного дослідження обґрунтовано висновок про формування багаторівневої системи міжнародно-правового регулювання ШІ, в якій універсальний рівень забезпечує легітимність та інклюзивність, міжрегіональний – консолідацію експертних підходів, а регіональний – інституційну ефективність та можливість переходу до юридично обов'язкових норм. Принципи, сформульовані на універсальному рівні, набувають конкретизації через рекомендації ОЕСР та стандарти МСЕ і втілюються у юридично обов'язкові приписи на регіональному рівні, що наразі реалізовано лише Радою Європи та ЄС. Стратегічним викликом залишається забезпечення взаємоузгодженості регіональних регуляторних режимів задля запобігання фрагментації глобального регуляторного простору та збереження балансу між стимулюванням інновацій і захистом прав людини.

Наукова новизна одержаних результатів полягає в тому, що дисертація є першим у вітчизняній науці міжнародного права комплексним дослідженням багаторівневої системи міжнародно-правового регулювання ШІ. Уперше запропоновано авторську дефініцію ШІ як комплексної соціотехнічної системи; систематизовано дванадцять принципів, що формують нормативну основу регулювання ШІ на універсальному рівні; обґрунтовано концепцію «ітеративного нормотворення»; виявлено п'ять тенденцій розвитку міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях. Удосконалено наукові положення щодо правового статусу ШІ та його кваліфікації як об'єкта регулювання *sui generis*, класифікації доктринальних підходів до визначення ШІ, функцій «м'якого права», ризикоорієнтованого підходу, переговорної динаміки формування юридично обов'язкових актів та взаємодоповнюваності європейських інструментів регулювання. Набули подальшого розвитку положення щодо ролі

резолюцій Генеральної Асамблеї ООН у формуванні міжнародно-правового регулювання ШІ, конвергенції етичних принципів регулювання ШІ навколо восьми ключових груп, система наукових знань про регіональні ініціативи у сфері міжнародно-правового регулювання ШІ.

Практичне значення одержаних результатів полягає в можливості їх використання у формуванні позиції України в ООН, Раді Європи та інших міжнародних організаціях, а також у нормотворчій діяльності, спрямованій на розбудову національної системи правового регулювання ШІ. Результати дослідження підтверджують стратегічну обґрунтованість зовнішньополітичного курсу України на запровадження європейських підходів до правового регулювання ШІ – як у межах Ради Європи, так і Європейського Союзу – та водночас засвідчують необхідність адаптації найкращих практик різних регіональних моделей, що поєднуюватиметься з реальними інституційними та економічними можливостями української держави в умовах повоєнної відбудови. Особливої актуальності ці завдання набувають з огляду на підписання Україною 15 травня 2025 року Рамкової конвенції Ради Європи про штучний інтелект та права людини, демократію та верховенство права, що ставить на порядок денний її ратифікацію та імплементацію, а також з огляду на необхідність адаптації національного законодавства до вимог Акту про ШІ в межах процесу європейської інтеграції.

Ключові слова: штучний інтелект, міжнародне право, правове регулювання, «м'яке право», «тверде право», права людини, сталий розвиток, міжнародні організації, ООН, ОЕСР, Рада Європи, Європейський Союз, право ЄС, єдиний цифровий ринок ЄС, Акт про ШІ

SUMMARY

Ihnatusha V. V. Formation and Development of the International Legal Regulation of Artificial Intelligence. – *Qualifying scientific work on the rights of a manuscript.*

Dissertation for the degree of Doctor of Philosophy in specialty 293 "International Law" (29 – International Relations). – V. M. Koretsky Institute of State and Law of the National Academy of Sciences of Ukraine, Kyiv, 2026.

The dissertation is a comprehensive study of the theoretical and applied foundations of the international legal regulation of artificial intelligence (hereinafter – AI) at the universal, interregional, and regional levels. The work examines the genesis and evolution of the international legal regulation of AI, defines the concept of AI, explores its legal nature and principal characteristics, and provides a general analysis of the specific features of its international legal regulation. Particular attention is devoted to the question of the legal status of AI: it is substantiated that contemporary international law does not confer upon AI an autonomous legal status, but rather qualifies it as an object of international legal regulation rather than a subject. The key instruments that determine the legal position of AI and govern its development, deployment, and use are analysed, with particular emphasis placed on the fact that contemporary international legal regulation of AI rests not only on formal legal prescriptions but also on ethical and value-based foundations – above all, on the human-centred approach and respect for human dignity, human rights, and fundamental freedoms. The study identifies the state, trends, and regularities in the formation of a multilevel system of international legal regulation of this technology, in particular the convergence of regulatory principles around a shared value-based core (human-centredness, transparency, accountability, non-discrimination) amid the simultaneous fragmentation of regulatory approaches, the extraterritorial influence of EU law through the "Brussels effect", and the proliferation of the risk-based model of regulation.

The relevance of the dissertation is determined by the rapid development of AI technologies and the unprecedented acceleration of law-making activity in this field, in particular the adoption in 2024 of the first international treaty on AI – the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225) – and the first horizontal EU act on AI to enter into force, Regulation (EU) 2024/1689 (the AI Act). Notwithstanding the growing number of instruments aimed at regulating AI, the doctrinal understanding of the legal status of AI and its regulation remains fragmentary, while a comprehensive study encompassing the entire multilevel system of the international legal regulation of AI – from the universal to the regional level – and its evolution from "soft" to "hard" law is absent in the Ukrainian science of international law.

The first chapter of the dissertation traces the historical trajectory of the development of the concept of artificial intelligence from the 1956 Dartmouth Conference, for the convening of which J. McCarthy proposed the very term "artificial intelligence", to the present stage of intensive law-making activity. It is established that the international legal regulation of AI developed in a non-linear manner: AI proceeded from a purely technical phenomenon, through ethical principles and instruments of "soft law", to legally binding norms. The study addresses the problem of the absence of a single definition of artificial intelligence in international law and legal doctrine, which generates uncertainty as to the boundaries of the subject matter of regulation and complicates the harmonisation of States' obligations. On the basis of an analysis of international legal instruments and legal doctrine, the author proposes an original classification of conceptual approaches to defining AI: technological, anthropocentric, functional, systemic, evolutionary, goal-oriented, and integrative. The author formulates an original definition of artificial intelligence as a complex sociotechnical system characterised by the functional capacity for inference (deriving conclusions from data), operational autonomy of varying degrees, adaptiveness, generative capacity, a cross-border character of application, dual-use potential, and the production of legally significant consequences for the sphere of human rights, the international legal order, and security. This definition embodies an integrative

approach, ensuring technological neutrality and suitability for use in international legal instruments of various levels.

A distinct place in the chapter is occupied by the question of the legal status of AI. It is substantiated that contemporary international law does not confer upon AI an autonomous legal status – in contrast to the State and the international organisation as generally recognised subjects of international law, and to the individual, for whom international law establishes specially defined status categories (refugee, stateless person, combatant). Instead, the legal nature of AI is established as an object of international legal regulation *sui generis*, characterised by a combination of material and immaterial components, of object-like and subject-like features, by dynamism and evolvability, as well as by a cross-border character of application, dual-use potential, autonomy of varying degrees, the capacity for self-improvement, and a substantial impact on international relations and their legal regulation. It is demonstrated that the fundamental differences between AI and traditional objects of international legal regulation necessitate the formation of a special legal regime that combines elements of various branches of international law.

It is established that the international legal regulation of AI is characterised by the fragmentation of regulatory approaches at all levels and, at the same time, by the steady convergence of principles around a shared value-based core: the human-centred approach, transparency and explainability, accountability and human oversight, safety and reliability, inclusiveness and non-discrimination, and the protection of privacy and data. It is emphasised that it is precisely these ethical and value-based orientations, rather than formal legal prescriptions alone, that form the value-normative foundation of contemporary AI regulation. The study reveals the specific features of the relationship between "soft" and "hard" law in the sphere of the international legal regulation of AI, in particular by distinguishing four functions of "soft law": the crystallisation of consensus, legal experimentation, legitimisation, and the preparation of the groundwork for legally binding regulation. At the same time, the conclusion is drawn as to the insufficiency of "soft law"

alone for the effective regulation of technologies entailing large-scale and potentially irreversible consequences.

In the second chapter of the dissertation, on the basis of a systematic analysis of 23 resolutions of the United Nations General Assembly adopted in the period 2015–2025, twelve principles are systematised that form the normative basis of the international legal regulation of AI at the universal level. An original chronology of the development of the UN General Assembly's law-making activity in the field of AI is proposed, and the following regularities are identified: the consolidation of an approach based on the "life cycle" of AI systems as the foundational regulatory framework; the standardisation of the triad "safe, secure and trustworthy AI systems"; and the positioning of the Global Digital Compact as the point of convergence of the legal regulation of AI at the UN level.

Also in the second chapter, it is demonstrated that the activity of the UN specialised agencies – UNESCO, ITU, WHO, WIPO, and ILO – forms a multidimensional system of AI regulation with a "gradation of normativity" in the form of "concentric circles": from instruments possessing the highest degree of political legitimacy but, at the same time, the highest degree of abstraction, to sectoral standards that become de facto binding owing to interoperability requirements. This system constitutes a polycentric institutional landscape without a clear hierarchy, in which instruments of different levels coexist and mutually reinforce one another.

In addition, in the second chapter, on the basis of the author's criteria, the appropriateness of distinguishing the interregional level as an autonomous subject of scholarly analysis is substantiated. It is established that the OECD acts as the central link of interregional cooperation in the sphere of AI regulation, with whose work such intergovernmental formats as the G7, the G20, and the GPAI coordinate their activity. The author proposes and substantiates the concept of "iterative law-making" as a model for the formation of "soft law" instruments in the sphere of the international legal regulation of AI, whereby an instrument functions as a "living document" that undergoes recurring cycles of adoption, testing, evaluation of effectiveness, and revision in response to technological change. The practical embodiment of this model is the updating of the OECD

Council Recommendation on Artificial Intelligence and the systematic revision of the ASEAN Guide on AI Governance and Ethics.

On the basis of the research conducted, key trends in the development of the international legal regulation of AI at the universal and interregional levels are identified and systematised: the acceleration of law-making activity, the transition from fragmentary to systemic regulation, the institutionalisation of regulation, the expansion of the subject matter and deepening of the detail of regulation, and the formation of the model of "iterative law-making". At the same time, persistent gaps are identified: the insufficient regulation of issues concerning the prevention of the adverse impact of AI on the environment, the liability of private-law subjects for the consequences of the functioning of AI systems, the challenges associated with generative AI, the impact of AI on the formation of a person's worldview and on their cognitive autonomy, the problems of artificial general intelligence (AGI) and long-term existential risks, as well as the issue of the military use of AI.

In the third chapter of the dissertation, it is established that it was precisely at the regional level that the transition began from the declaration of principles of the international legal regulation of AI to the formulation of legally binding norms – a transition thus far implemented only by the European Union, through Regulation (EU) 2024/1689 (the AI Act), and by the Council of Europe, through the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225). A periodisation of the formation of the EU legal framework in the field of AI regulation is proposed, and its systemic regularities are identified: the dualism of market-based and value-based justifications of regulation, the gradual weakening of conformity-assurance mechanisms, and the reactive character of regulation. It is substantiated that the predominance of the risk-based approach in contemporary international legal regulation of AI reflects the application of the principle of proportionality and constitutes a synthesis of the precautionary approach with the aspiration to stimulate innovation, thereby enabling a balance to be struck between the protection of human rights and the interests of technological development. The fullest embodiment of this approach is the AI Act, which

introduces a four-tier classification of AI systems by degree of risk (unacceptable, high, limited, and minimal) and differentiates the scope of regulatory requirements according to the level of potential harm to human rights, safety, and fundamental values.

On the basis of a comparative analysis of the zero draft and the final text of the Council of Europe Framework Convention, a trend is established towards the systematic narrowing of the scope of States' obligations, which occurred in the course of intergovernmental negotiations. This provides grounds for the conclusion as to the existence of the following general regularity: at the negotiation stage, the human-rights-protective potential of legally binding instruments on AI regulation is consistently diminished under the influence of both States and private-law subjects. At the same time, the complementarity of the two European instruments of differing legal nature is demonstrated – the Framework Convention, which implements a human-rights-based approach, and the AI Act, which implements a risk-based approach. The study also examines the adaptive potential of the Convention for the Protection of Human Rights and Fundamental Freedoms of 1950 (the European Convention on Human Rights) with respect to AI-related matters, as a result of which it is established that the case law of the European Court of Human Rights is gradually forming a body of precedent that adapts to digital realities; however, this case law is predominantly reactive in character and cannot replace the proactive legal regulation of AI effected through the Framework Convention and the AI Act.

A comparative analysis of the approaches to the legal regulation of AI elaborated by regional organisations and associations of States of other regions – ASEAN, APEC, the SCO, the African Union, the OAS, the Cooperation Council for the Arab States of the Gulf, and ALECSO – made it possible to identify the following key regularities: a "spectrum of legal bindingness", under which "hard law" remains a phenomenon peculiar exclusively to the European region; an "axiological consensus amid regulatory diversity", under which each region contributes to the global discourse its own values and regulatory approaches – from the philosophy of Ubuntu in the instruments of the African Union and Islamic ethical concepts in the initiatives of the Gulf States to the consensus-based model of ASEAN; the

differentiated character of the "Brussels effect"; the "paradox of global AI regulation", under which the regions with the greatest needs for the protection of human rights possess the least institutional capacity; and the formation of a "regulatory polylogue" among different regions.

As a result of the research conducted, the conclusion is substantiated as to the formation of a multilevel system of the international legal regulation of AI, in which the universal level ensures legitimacy and inclusiveness, the interregional level ensures the consolidation of expert approaches, and the regional level ensures institutional effectiveness and the possibility of transition to legally binding norms. The principles formulated at the universal level acquire concretisation through OECD recommendations and ITU standards and are embodied in legally binding prescriptions at the regional level, which thus far has been realised only by the Council of Europe and the EU. A strategic challenge remains the assurance of the mutual coherence of regional regulatory regimes in order to prevent the fragmentation of the global regulatory space and to preserve the balance between the stimulation of innovation and the protection of human rights.

The scientific novelty of the results obtained consists in the fact that the dissertation is the first comprehensive study, in the Ukrainian science of international law, of the multilevel system of the international legal regulation of AI. For the first time, an original definition of AI as a complex sociotechnical system is proposed; twelve principles forming the normative basis of AI regulation at the universal level are systematised; the concept of "iterative law-making" is substantiated; and five trends in the development of the international legal regulation of AI at the universal and interregional levels are identified. The scholarly propositions concerning the legal status of AI and its qualification as an object of regulation *sui generis*, the classification of doctrinal approaches to defining AI, the functions of "soft law", the risk-based approach, the negotiation dynamics of the formation of legally binding instruments, and the complementarity of the European regulatory instruments have been improved. Further development has been given to the propositions concerning the role of UN General Assembly resolutions in the formation of the international legal regulation of AI, the convergence of ethical principles of AI

regulation around eight key clusters, and the body of scholarly knowledge concerning regional initiatives in the sphere of the international legal regulation of AI.

The practical significance of the results obtained consists in the possibility of their use in the formation of Ukraine's position within the UN, the Council of Europe, and other international organisations, as well as in law-making activity aimed at developing a national system of legal regulation of AI. The results of the research confirm the strategic soundness of Ukraine's foreign-policy course towards the adoption of European approaches to the legal regulation of AI – both within the framework of the Council of Europe and that of the European Union – and at the same time attest to the necessity of adapting the best practices of various regional models in a manner consonant with the real institutional and economic capacities of the Ukrainian State in the conditions of post-war reconstruction. These tasks acquire particular relevance in light of Ukraine's signing, on 15 May 2025, of the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, which places on the agenda its ratification and implementation, as well as in light of the need to adapt national legislation to the requirements of the AI Act within the process of European integration.

Keywords: artificial intelligence, international law, legal regulation, soft law, hard law, human rights, sustainable development, international organisations, UN, OECD, Council of Europe, European Union, EU law, EU Digital Single Market, AI Act.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні наукові результати дисертації:

1. Ігнатуша В. В. Розвиток міжнародно-правового регулювання штучного інтелекту в рамках ООН. Аналітично-порівняльне правознавство. 2025. № 3, ч. 3. С. 287–296. DOI: 10.24144/2788-6018.2025.03.3.45. URL:

<https://journal-app.uzhnu.edu.ua/article/view/335016>

2. Ігнатуша В. В. Автономні системи з ШІ у збройних конфліктах та правове забезпечення дотримання принципу розрізнення. Науковий вісник Ужгородського національного університету. Серія «Право». 2025. Вип. 90, ч. 5. С. 151–161. DOI: 10.24144/2307-3322.2025.90.5.18. URL:

<https://visnyk-pravo.uzhnu.edu.ua/article/view/340844>

3. Ігнатуша В. В. Міжнародно-правові механізми забезпечення контролю людини над розвитком систем штучного інтелекту. Український часопис міжнародного права. 2025. № 2. С. 67–71. DOI: 10.36952/ujil.2025.2.67-71. URL:

<https://jusintergentes.com.ua/images/2025/2/67-71.pdf>

4. Ігнатуша В. В. Правове регулювання принципів прозорості та підзвітності систем ШІ в ЄС: сучасний стан та перспективи розвитку. Юридичний науковий електронний журнал. 2026. № 1. С. 265–272. DOI: 10.32782/2524-0374/2026-1/58. URL: https://www.lsej.org.ua/1_2026/60.pdf

Наукові праці, які засвідчують апробацію матеріалів дисертації:

5. Ігнатуша В. В. Ризики використання штучного інтелекту в оборонній сфері та способи їх подолання в рамках ЄС та НАТО. Інтеграція України до Європейського Союзу: виклики в умовах воєнного стану. Пам'яті професора Павла Заворотька та професора Володимира Денисова : зб. матеріалів і тез доп. міжнар. наук.-практ. конф. (м. Київ, 4–6 листопада 2024 р.) / за наук. ред. проф. Л. Г. Фалалєєвої. Київ :

Вид-во «Людмила», 2025. С. 299–303. URL: <https://idpnan.org/integration-into-european-union>

6. Ihnatusha V. Classification of AI systems in international legal regulation: a comparative approach. Science in the Context of Modern Challenges and Prospects : proceedings of the International Scientific Conference (May 21, 2025). Utrecht : Bookmundo, 2025. P. 227–230. DOI: 10.64076/iedc250521.

7. Ігнатуша В. В. Нормативне закріплення принципу контролю людини над системами ШІ у міжнародно-правових актах. Глобальні виклики та інновації: шляхи розвитку сучасної науки : матеріали X Міжнар. наук. конф. (м. Одеса, 16 січня 2026 р.). Одеса, 2026. С. 235–238. DOI: 10.62731/mcnd-16.01.2026.

8. Ігнатуша В. В. Превентивні інструменти забезпечення контролю людини над системами ШІ у міжнародному праві. The Current State of Development of World Science: Characteristics and Features : proceedings of the X International Scientific and Theoretical Conference (Lisbon, January 23, 2026). Lisbon, 2026. P. 150–153. DOI: 10.36074/scientia-23.01.2026.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	4
ВСТУП	6
РОЗДІЛ 1. ХАРАКТЕРИСТИКА ШТУЧНОГО ІНТЕЛЕКТУ ЯК ОБ'ЄКТА МІЖНАРОДНО-ПРАВОВОГО РЕГУЛЮВАННЯ	21
1.1. Генеза та еволюція міжнародно-правового регулювання штучного інтелекту	21
1.2. Поняття та основні властивості штучного інтелекту як об'єкта міжнародно- правового регулювання	34
1.3. Загальний аналіз специфіки міжнародно-правового регулювання штучного інтелекту	60
Висновки до Розділу 1	76
РОЗДІЛ 2. МІЖНАРОДНО-ПРАВОВЕ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ НА УНІВЕРСАЛЬНОМУ ТА МІЖРЕГІОНАЛЬНОМУ РІВНЯХ	82
2.1. Напрямки роботи ООН щодо правового регулювання ШІ в інтересах забезпечення сталого розвитку	82
2.2. Ініціативи спеціалізованих установ ООН у сфері регулювання штучного інтелекту	104
2.3. Роль ОЕСР та міжрегіональних ініціатив у формуванні міжнародних стандартів застосування штучного інтелекту	124
Висновки до Розділу 2	144
РОЗДІЛ 3. МІЖНАРОДНО-ПРАВОВЕ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ НА РЕГІОНАЛЬНОМУ РІВНІ	150

3.1. Формування правової бази Європейського Союзу у сфері регулювання штучного інтелекту	150
3.2. Діяльність Ради Європи щодо захисту прав людини в умовах розвитку технологій штучного інтелекту	168
3.3. Інші регіональні ініціативи у сфері міжнародно-правового регулювання штучного інтелекту: порівняльний аналіз	189
Висновки до Розділу 3	211
ВИСНОВКИ	215
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	225
ДОДАТКИ	269

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

АЛЕКСО	Арабська організація з питань освіти, культури і науки
АС	Африканський Союз
АСЕАН	Асоціація держав Південно-Східної Азії
АТЕС	Азійсько-Тихоокеанське економічне співробітництво
ВВУЮ	Всесвітня зустріч на вищому рівні з питань інформаційного суспільства
ВОІВ	Всесвітня організація інтелектуальної власності
ВООЗ	Всесвітня організація охорони здоров'я
ГА ООН	Генеральна Асамблея Організації Об'єднаних Націй
ГЦД	Глобальний цифровий договір
ДФЄС	Договір про функціонування Європейського Союзу
ЕКЛАК	Економічна комісія ООН для Латинської Америки та Карибського басейну
ЕКОСОП	Економічна і Соціальна Рада ООН
ЄКЕП	Європейська комісія з ефективності правосуддя
ЄКПЛ	Конвенція про захист прав людини і основоположних свобод
ЄС	Європейський Союз
ЄСПЛ	Європейський суд з прав людини
ЛАД	Ліга арабських держав
МАГАТЕ	Міжнародне агентство з атомної енергії
МЕРКОСУР	Південноамериканський спільний ринок
МОП	Міжнародна організація праці
МСЕ	Міжнародний союз електрозв'язку
ОАД	Організація американських держав
ОБСЄ	Організація з безпеки і співробітництва в Європі

ОЕСР	Організація економічного співробітництва та розвитку
ООН	Організація Об'єднаних Націй
ПАРЄ	Парламентська Асамблея Ради Європи
ПРООН	Програма розвитку ООН
РЄ	Рада Європи
РСАДПЗ	Рада співробітництва арабських держав Перської затоки
ЦСР	Цілі сталого розвитку
ШІ	штучний інтелект
ШОС	Шанхайська організація співробітництва
ЮНЕСКО	Організація Об'єднаних Націй з питань освіти, науки і культури
AGI	штучний загальний інтелект
AI HLEG	Експертна група високого рівня зі штучного інтелекту
AIGO	Робоча група ОЕСР з управління ШІ
САНАІ	Спеціальний міжурядовий комітет зі штучного інтелекту
СОМСУТ	Міжамериканський комітет з науки та технологій ОАД
DSA	Акт про цифрові послуги
ЕАС	Східноафриканське Співтовариство
ЕСОВАС	Економічне співтовариство держав Західної Африки
GDPR	Загальний регламент про захист даних
GPAI	Global Partnership on Artificial Intelligence
GPSR	Загальний регламент про безпеку продукції
LAWS	летальні автономні системи зброї
LLM	великі мовні моделі

ВСТУП

Обґрунтування вибору теми дослідження. Стрімкий розвиток технологій штучного інтелекту (далі – ШІ) став одним із визначальних чинників трансформації міжнародних відносин та їх правового регулювання у XXI столітті. Від Дартмутської конференції 1956 року, для проведення якої Дж. Маккарті запропонував сам термін «штучний інтелект», до сучасного періоду, за якого системи ШІ проникають у всі сфери суспільного життя – від охорони здоров'я та правосуддя до воєнної справи та міжнародної безпеки, – минуло понад сім десятиліть. Однак саме останнє десятиліття характеризується безпрецедентним прискоренням як технологічного розвитку, так і спроб його міжнародно-правового регулювання, що зумовлює актуальність комплексного наукового дослідження цього процесу.

Усвідомлення трансформаційного потенціалу ШІ для міжнародних відносин стало поворотним моментом у переході від суто технічного до правового дискурсу. Можливості, стрімкий розвиток і широкий спектр застосування технологій ШІ породили якісно нові суспільні відносини – щодо розроблення, розгортання та використання відповідних систем, – які вимагають формування системного міжнародно-правового регулювання. Спроби концептуалізувати вплив ШІ на міжнародне право знайшли відображення, зокрема, у дослідженні М.М. Мааса, який виокремлює три типи такого впливу: розвиток нових норм для регулювання ШІ (legal development), заміщення чинних правил через автоматизацію процесів (legal displacement) та руйнування традиційних правових інститутів через фундаментальну зміну умов їхнього функціонування (legal destruction) [2, с. 38]. Ця типологія наочно ілюструє масштаб викликів, які штучний інтелект створює для сучасного міжнародного права. Водночас жоден із цих типів впливу не передбачає визнання штучного інтелекту самостійним носієм яких-небудь прав чи обов'язків; міжнародне право реагує на феномен ШІ не шляхом конструювання для нього спеціального

правового статусу, а шляхом формування правового режиму, у межах якого ШІ розглядається як об'єкт регулювання, а не суб'єкт.

Міжнародно-правове регулювання ШІ формується на кількох рівнях одночасно: універсальному (в рамках ООН та її спеціалізованих установ), міжрегіональному (ОЕСР, G7, G20, GRAI) та регіональному (Європейський Союз, Рада Європи, АСЕАН, Африканський Союз, ОАД та інші). У сфері регулювання ШІ постає особлива система, що поєднує «тверде» та «м'яке право» і виникає зі складних взаємодій численних джерел права та учасників відносин: держав, міжнародних організацій, суб'єктів приватного права, громадянського суспільства та наукових установ. Ця багаторівнева та поліцентрична архітектура регулювання, що охоплює широкий спектр актів різної юридичної сили – від декларацій та рекомендацій до перших юридично обов'язкових актів: Рамкової конвенції Ради Європи про штучний інтелект та права людини, демократію та верховенство права 2024 р. (CETS № 225) як першого міжнародного договору щодо ШІ та Регламенту (ЄС) 2024/1689 (Акту про ШІ) як першого горизонтального акта щодо ШІ, що набув чинності, – потребує системного наукового осмислення.

Актуальність теми дисертаційного дослідження підсилюється тим, що процес міжнародно-правового регулювання ШІ зазнав кардинального прискорення саме протягом останніх років. Якщо перші систематизовані спроби формування правової бази розпочалися лише в середині 2010-х років із Резолюції Європейського парламенту 2017 року з рекомендаціями щодо цивільно-правових норм у сфері робототехніки та Рекомендації Ради ОЕСР щодо штучного інтелекту 2019 року, то вже у 2024 році ухвалено низку ключових актів: першу в історії резолюцію Генеральної Асамблеї ООН, спеціально присвячену штучному інтелекту (A/RES/78/265); Глобальний цифровий договір як складову Пакту про майбутнє (A/RES/79/1); Рамкову конвенцію Ради Європи (CETS № 225); Регламент (ЄС) 2024/1689 (Акт про ШІ). Одним з пріоритетних напрямів розвитку міжнародної нормотворчості на сучасному етапі є врахування досягнень новітніх технологій,

пов'язаних зі штучним інтелектом, що додатково підкреслює своєчасність обраної тематики.

Разом з тим, незважаючи на стрімке зростання кількості міжнародно-правових актів, спрямованих на регулювання ШІ, доктринальне осмислення цього процесу залишається фрагментарним. Зарубіжні дослідження здебільшого фокусуються на окремих рівнях регулювання, конкретних міжнародних організаціях чи міжнародно-правових актах, не забезпечуючи цілісного бачення формування міжнародно-правового регулювання ШІ. Теоретичною базою дисертаційного дослідження є, зокрема, науковий доробок М.М. Мааса (M.M. Maas), який запропонував типологію впливу ШІ на міжнародне право, М. Шерера (M. Scherer), який досліджував проблематику правового визначення ШІ та регуляторних викликів, М. Чінена (M. Chinen), А. Гузмана (A. Guzman) та Т. Меєра (T. Meyer), які розробляли теоретичні засади міжнародного «м'якого права» та регулювання нових технологій. Інституційну архітектуру та нормативні інструменти регулювання ШІ досліджували Й. Таллберг (J. Tallberg), Г. Робертс (H. Roberts), М. Таддео (M. Taddeo) та Л. Флоріді (L. Floridi), а поліцентричний характер міжнародно-правового регулювання ШІ аналізували П. Сайхон (P. Cihon), Л. Кемп (L. Kemp) та Л. Шмітт (L. Schmitt). Проблематику конвергенції етичних принципів регулювання ШІ висвітлювали Дж. Фелд (J. Fjeld) та Н. Корреа (N. Correia), а А. Каушкік (A. Kaushik) досліджував нормативний потенціал Рекомендації ЮНЕСКО з етики ШІ. Значний внесок у дослідження регуляторних інструментів європейського регіону здійснили М.А. Пресно Лінера (M.A. Presno Linera) та А. Мьовезе (A. Meuwese), які провели порівняльний аналіз Акту про ШІ та Рамкової конвенції Ради Європи, М. Алмада (M. Almada) та А. Раду (A. Radu), які досліджували екстратериторіальний вплив Акту про ШІ, а також А. Бредфорд (A. Bradford), яка обґрунтувала концепцію «Брюссельського ефекту» у контексті глобального поширення регуляторних стандартів. Підходи до регулювання ШІ за межами Європи аналізувалися у працях Б.А. Путри (B.A. Putra) щодо АСЕАН, Ц. Сюй (C. Xu), Т. Лі (T. Lee) та Г. Гоггіна (G. Goggin) щодо Азійсько-Тихоокеанського регіону, У.Ч. Аджузіогу (U.C. Ajuzieogu)

та М.Б. Юнвера (M.B. Unver) щодо Африканського Союзу, а також С. Фельдштейна (S. Feldstein), Е. Кляйн (E. Klein), С. Патріка (S. Patrick) та інших зарубіжних авторів.

У вітчизняній юридичній науці, у тому числі в науці міжнародного права, проблематика міжнародно-правового регулювання ШІ досліджена недостатньо і має переважно фрагментарний характер. Серед вітчизняних дослідників, які зверталися до окремих аспектів цієї проблематики, слід відзначити Н.М. Оніщенко, яка досліджувала вплив новітніх технологій на розвиток нормотворчості; О.В. Костенка, який розкривав правові виміри новітніх цифрових технологій, зокрема обґрунтовував тезу про ненейтральність регуляторного «вичікування» у їх правовому регулюванні; М.І. Суржинського, який аналізував виклики штучного інтелекту для національної правотворчості та обґрунтовував необхідність адаптивного регулювання у цій сфері; Н.О. Попову, яка досліджувала вплив ШІ на міжнародний правопорядок та обґрунтовувала необхідність формування відповідної законодавчої бази; М.О. Медведєву, яка аналізувала вплив технологій ШІ на систему сучасного міжнародного права та його окремі галузі, а також міжнародно-правове регулювання ШІ в контексті охорони навколишнього середовища; О.А. Баранова, який аналізував концептуальні засади визначення штучного інтелекту через призму еквівалентних характерних властивостей; Н.В. Камінську, яка аналізувала трансформацію міжнародного та наднаціонального права під впливом ШІ; І.М. Берназюка, який розглядав виклики ШІ для Конвенції про захист прав людини і основоположних свобод 1950 року (Європейської конвенції з прав людини); А. Гачкевича, який обґрунтовував виникнення права людини на безпечне використання штучного інтелекту; О.Е. Радутного, який порушував питання кримінальної відповідальності штучного інтелекту; Д.М. Белова та М.В. Белову, які досліджували інтеграцію технологій ШІ у правову систему. Деякі аспекти правового регулювання ШІ розглядалися також у працях Г. Машлій, О. Мосій та М. Пельчер, О. Кожухар та інших вітчизняних авторів. Водночас наявні публікації зосереджені переважно на вузьких аспектах регулювання ШІ – конституційно-правових, кримінально-правових, цивільно-правових чи правозахисних, – тоді як комплексне дослідження,

яке б охоплювало всю багаторівневу систему міжнародно-правового регулювання штучного інтелекту – від універсального до регіонального рівня, від «м'якого» до «твердого» права, від декларативних принципів до юридично обов'язкових норм, – у вітчизняній науці відсутнє. Саме тому існує нагальна потреба у комплексному вивченні всього спектру питань, пов'язаних із формуванням міжнародно-правового регулювання ШІ, що підкреслює актуальність обраної теми дисертації.

Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертацію виконано у рамках планової науково-дослідної теми відділу міжнародного права та права Європейського Союзу Інституту держави і права імені В. М. Корецького НАН України «Принцип незастосування сили в сучасних умовах функціонування колективної безпеки» (номер державної реєстрації РК 0120U105286 та «Адаптація законодавства України до права Європейського Союзу в умовах повоєнного відновлення» (номер державної реєстрації РК 0123V104299).

Мета і завдання дослідження. Метою дисертаційного дослідження є комплексна теоретична розробка проблематики міжнародно-правового регулювання штучного інтелекту: розкриття його сутності, з'ясування сучасного стану та визначення тенденцій розвитку на універсальному, міжрегіональному та регіональному рівнях. Концептуалізація теоретичних засад та перспектив розвитку міжнародно-правового регулювання ШІ проводилася з огляду на пріоритетні напрями його еволюції, а також різні аспекти діяльності міжнародних організацій у цій сфері, враховуючи специфіку класичного міжнародного права та динамічний характер розвитку технологій штучного інтелекту.

Для досягнення зазначеної мети вирішувалися такі дослідницькі завдання: дослідити генезу та еволюцію міжнародно-правового регулювання штучного інтелекту; визначити поняття та основні характеристики штучного інтелекту як об'єкта міжнародно-правового регулювання, з'ясувавши його правову природу; виявити та проаналізувати специфіку міжнародно-правового регулювання штучного інтелекту, включаючи особливості співвідношення «м'якого» та «твердого» права, фрагментацію та конвергенцію регуляторних підходів; проаналізувати нормотворчу

діяльність ООН щодо регулювання ШІ, насамперед резолюції Генеральної Асамблеї ООН, ухвалені в контексті забезпечення сталого розвитку та міжнародної безпеки; дослідити ініціативи спеціалізованих установ ООН (ЮНЕСКО, МСЕ, ВООЗ, ВОІВ, МОП) у сфері регулювання штучного інтелекту; визначити роль ОЕСР та пов'язаних з нею міжрегіональних форматів (G7, G20, GPAI) у формуванні міжнародних стандартів регулювання ШІ; дослідити формування правової бази Європейського Союзу у сфері регулювання штучного інтелекту; проаналізувати діяльність Ради Європи щодо захисту прав людини в умовах розвитку технологій штучного інтелекту; здійснити порівняльний аналіз ініціатив АСЕАН, Африканського Союзу, ОАД, Ради співробітництва арабських держав Перської затоки, ШОС, АТЕС та інших регіональних організацій у сфері регулювання ШІ.

Об'єкт дослідження становлять міжнародні відносини у сфері розроблення, розгортання та використання штучного інтелекту.

Предметом дослідження є міжнародно-правове регулювання штучного інтелекту, його становлення та розвиток.

Методи дослідження. Дисертаційне дослідження виконано з використанням філософських, загальнонаукових і спеціально-наукових методів, що дало можливість всебічно висвітлити формування та еволюцію міжнародно-правового регулювання ШІ. В основу дослідження покладено методологічний плюралізм, завдяки якому розкрито вплив діяльності міжнародних організацій на розвиток регулювання ШІ. До філософських методів пізнання належать насамперед діалектичний, герменевтичний та аксіологічний методи. *Діалектичний метод* було застосовано під час аналізу еволюції міжнародно-правового регулювання ШІ, яка становить єдність і боротьбу протилежностей: фрагментації регуляторних підходів та конвергенції принципів; стимулювання технологічних інновацій та захисту прав людини від ризиків, пов'язаних із застосуванням ШІ. Водночас діалектичний метод забезпечив висвітлення теми дисертації у взаємозв'язку юридичних питань з технологічними, геополітичними, економічними та етичними аспектами. *Герменевтичний метод* використовувався для інтерпретації приписів міжнародно-

правових актів, зокрема резолюцій, рекомендацій, конвенцій та регламентів – актів різної правової природи, на основі яких формується міжнародно-правове регулювання ШІ. Застосування цього методу було зумовлене тим, що значна частина аналізованих актів оперує оціночними поняттями та відкритими формулюваннями – такими як «безпечні, захищені та надійні системи ШІ», «людиноцентричний підхід», «пояснюваність», «належний людський нагляд», – зміст яких не може бути з'ясований виключно на підставі буквального тексту. Герменевтичний метод дозволив розкрити нормативний зміст цих понять через їх телеологічну інтерпретацію (з'ясування цілей, що переслідувалися під час ухвалення акта), системне тлумачення (встановлення зв'язків між приписами різних актів та різних рівнів регулювання), а також контекстуальний аналіз з урахуванням підготовчих матеріалів та інституційної практики відповідних міжнародних організацій. Завдяки *аксіологічному методу* проаналізовано ціннісні засади, покладені в основу міжнародно-правового регулювання ШІ, зокрема повагу до людської гідності, захист прав людини, демократичні цінності та верховенство права.

Серед загальнонаукових методів, що використовувались у дисертаційній роботі, – формально-логічний, системний та синергетичний методи. За допомогою *формально-логічного методу* розкрито логічні зв'язки між положеннями міжнародно-правових актів різного рівня та різної юридичної сили щодо регулювання ШІ. *Системний метод* дозволив дослідити міжнародно-правове регулювання ШІ як цілісну багаторівневу систему, в якій універсальний рівень забезпечує легітимність та інклюзивність, міжрегіональний – консолідацію експертних підходів, а регіональний – інституційну ефективність та можливість переходу від декларування принципів до юридично обов'язкових норм. За допомогою *синергетичного методу* досліджено нелінійну динаміку формування міжнародно-правового регулювання ШІ, зокрема стрибкоподібне прискорення нормотворчої активності протягом 2023–2025 років, а також біфуркаційні точки розвитку регулювання, в яких ухвалення ключових актів спричинило якісні зміни на всіх досліджуваних рівнях.

До спеціально-наукових методів, які було використано в дисертаційному дослідженні, належать передовсім формально-юридичний, порівняльно-правовий та історико-правовий методи. Застосування *формально-юридичного методу* сприяло аналізу міжнародно-правових актів, що регулюють сферу ШІ, з'ясуванню їхньої правової природи, юридичної сили та механізмів імплементації. *Порівняльно-правовий метод* дав можливість зіставити регуляторні підходи різних міжнародних організацій та регіональних об'єднань – ЄС, Ради Європи, АСЕАН, Африканського Союзу, ОАД, Ради співробітництва арабських держав Перської затоки та інших, виявити закономірності та розбіжності у міжнародно-правовому регулюванні ШІ на різних рівнях. Завдяки застосуванню *історико-правового методу* проаналізовано різні етапи становлення та розвитку міжнародно-правового регулювання ШІ у хронологічній послідовності та з урахуванням специфіки кожного з них.

Наукова новизна отриманих результатів полягає в тому, що дисертація є комплексним науковим дослідженням, яке заповнює в українській науці міжнародного права прогалину у висвітленні теоретико-прикладних засад міжнародно-правового регулювання штучного інтелекту на універсальному, міжрегіональному та регіональному рівнях. Наукову новизну дисертаційної роботи та особистий внесок автора у дослідження цієї проблематики становлять такі основні положення:

уперше:

запропоновано авторську дефініцію штучного інтелекту як комплексної соціотехнічної системи, що характеризується функціональною здатністю до інференції (виведення висновків з даних), операційною автономністю різного ступеня, адаптивністю, генеративною спроможністю, транскордонним характером застосування, потенціалом подвійного використання та створенням юридично значущих наслідків для сфери прав людини, міжнародного правопорядку та безпеки; дефініція втілює інтегративний підхід, поєднуючи технологічні, функціональні, системні, антропоцентричні, еволюційні та правові елементи, і забезпечує

технологічну нейтральність визначення та його придатність для використання у міжнародно-правових актах різного рівня;

систематизовано – на підставі системного аналізу двадцяти трьох резолюцій Генеральної Асамблеї ООН 2015–2025 років – авторський перелік із дванадцяти принципів, безпосередньо закріплених у цих резолюціях і таких, що формують нормативну основу міжнародно-правового регулювання ШІ на універсальному рівні: людиноцентричний підхід; безпека, захищеність та надійність; прозорість та пояснюваність; підзвітність та людський нагляд; інклюзивність та недискримінація; міжнародне співробітництво; подолання цифрового розриву; сприяння сталому розвитку; захист приватності та даних; підхід за участю багатьох зацікавлених сторін; обережність (запобіжний підхід); застосовність міжнародного права;

запропоновано та обґрунтовано концепцію «ітеративного нормотворення» як моделі формування актів «м'якого права» у сфері міжнародно-правового регулювання ШІ, за якої нормативний акт не розглядається як остаточна фіксація норм, а функціонує як «живий документ», що проходить повторювані цикли прийняття, апробації, оцінки ефективності та перегляду у відповідь на технологічні зміни. Модель ітеративного нормотворення відрізняється від класичної парадигми міжнародного нормотворення за трьома ключовими ознаками: нормативна незавершеність як свідомий вибір, скорочення тривалості нормотворчих циклів, а також інкрементальність змін із збереженням ціннісного ядра акта;

виявлено та систематизовано п'ять ключових тенденцій розвитку міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях: прискорення нормотворчої діяльності; перехід від фрагментарного до системного регулювання; інституціоналізація регулювання через створення постійних структур; розширення предмета та поглиблення деталізації регулювання – від загальних декларацій до регламентації конкретних сфер застосування ШІ; поступове формування моделі ітеративного нормотворення;

удосконалено:

наукові положення щодо правового статусу штучного інтелекту в міжнародному праві: обґрунтовано, що чинне міжнародне право не наділяє штучний інтелект самостійним правовим статусом, розглядаючи його як об'єкт регулювання, – на відміну від держави та міжнародної організації як загальноновизнаних суб'єктів міжнародного права, а також індивіда, для якого міжнародне право встановлює спеціально визначені статусні категорії (біженця, особи без громадянства, комбатанта); розкрито правову природу штучного інтелекту як об'єкта міжнародно-правового регулювання *sui generis*, якій притаманне поєднання матеріальних і нематеріальних складових, об'єктних і суб'єктних рис, динамічності та еволюційності; доведено, що ця природа разом із транскордонним характером застосування ШІ та його суттєвим впливом на міжнародні відносини та їх правове регулювання зумовлює необхідність формування спеціального правового режиму, який інтегрує елементи різних галузей міжнародного права;

наукові положення щодо доктринальних підходів до визначення поняття штучного інтелекту в міжнародному праві – шляхом авторської класифікації семи концептуальних підходів, виокремлених у міжнародно-правових актах та науковій доктрині (технологічного, антропоцентричного, функціонального, системного, еволюційного, цілеорієнтованого та інтегративного);

наукові позиції щодо співвідношення «м'якого» та «твердого» права у сфері регулювання ШІ, зокрема виокремлено чотири функції «м'якого права»: кристалізація консенсусу, правовий експеримент, легітимізація та підготовка підґрунтя для юридично обов'язкового регулювання, а також обґрунтовано недостатність самого лише «м'якого права» для ефективного регулювання технологій з масштабними та потенційно необоротними наслідками;

концептуальні засади ризикоорієнтованого підходу до регулювання ШІ як прояву принципу пропорційності в міжнародному праві та синтезу запобіжного підходу (принципу обережності) з прагненням до стимулювання інновацій, що дозволяє забезпечувати баланс між захистом прав людини та інтересами технологічного розвитку;

наукові уявлення про нормотворчу діяльність Генеральної Асамблеї ООН у сфері ШІ: на матеріалі двадцяти трьох резолюцій 2015–2025 років виявлено три ключові закономірності – систематизацію підходу на основі «життєвого циклу» систем штучного інтелекту як базової регуляторної основи, стандартизацію тріади «безпечні, захищені та надійні системи штучного інтелекту» та позиціонування Глобального цифрового договору як точки конвергенції правового регулювання ШІ на рівні ООН;

наукові положення про функціональну спеціалізацію спеціалізованих установ ООН (ЮНЕСКО, МСЕ, ВООЗ, ВОІВ, МОП) у сфері регулювання ШІ: обґрунтовано, що їхня діяльність у межах функціональних мандатів формує багатовимірну систему секторального регулювання з «градацією нормативності» – від етичного фундаменту Рекомендації ЮНЕСКО, на який спираються ініціативи інших установ, через технічні стандарти МСЕ, що набувають де-факто обов'язковості завдяки вимогам інтероперабельності, та галузеві керівництва ВООЗ – до аналітичних напрацювань ВОІВ і МОП, які виконують функцію попереднього дослідження;

наукові положення щодо переговорної динаміки формування юридично обов'язкових актів у сфері регулювання ШІ: на підставі порівняльного аналізу нульового проєкту та фінального тексту Рамкової конвенції Ради Європи про ШІ, а також еволюції проєкту Акту про ШІ від пропозиції Європейської комісії до ухваленого тексту встановлено загальну закономірність систематичного звуження обсягу зобов'язань на етапі переговорів під впливом як держав, так і суб'єктів приватного права;

положення про взаємодоповнюваність двох європейських інструментів регулювання ШІ різної правової природи – Рамкової конвенції Ради Європи як міжнародного договору, що реалізує правозахисний підхід, та Регламенту (ЄС) 2024/1689 (Акту про ШІ) як інструмента права ЄС, що реалізує ризикоорієнтований підхід, – поєднання яких формує найбільш комплексну з усіх досліджених регіональних регуляторних систем;

наукові положення щодо регіонального рівня міжнародно-правового регулювання ШІ: за результатами порівняльного аналізу європейських та позаєвропейських регіональних ініціатив ідентифіковано п'ять закономірностей – «спектр юридичної обов'язковості», в якому «тверде право» залишається феноменом, притаманним виключно європейському регіону; «аксіологічний консенсус при регуляторному розмаїтті»; диференційований характер «Брюссельського ефекту»; «парадокс глобального регулювання ШІ», за якого регіони з найбільшими потребами у захисті прав людини мають найменші інституційні спроможності; формування «регуляторного полілогу» між різними регіонами;

набули подальшого розвитку:

наукові позиції щодо ролі резолюцій Генеральної Асамблеї ООН у формуванні міжнародно-правового регулювання ШІ, зокрема їхнього значення для кристалізації міжнародного консенсусу щодо базових принципів регулювання штучного інтелекту, утвердження нормативних очікувань та стандартів поведінки держав, а також створення підґрунтя для майбутніх договірних зобов'язань;

положення наукової доктрини про конвергенцію етичних принципів регулювання ШІ навколо восьми ключових груп (прозорість та пояснюваність, підзвітність та відповідальність, справедливість та недискримінація, приватність та захист даних, безпека та захищеність, людський контроль, професійна відповідальність, просування людських цінностей): на відміну від виведеного в роботі переліку принципів резолюцій ГА ООН, ці групи є узагальненням етичних орієнтирів усього масиву актів у сфері регулювання ШІ; на матеріалі актів універсального та міжрегіонального рівнів цю конвергенцію підтверджено та конкретизовано, а також доведено, що вона створює основу для формування універсальних норм міжнародного права у сфері регулювання ШІ;

концепція багаторівневої системи міжнародно-правового регулювання ШІ, яка характеризується асиметрією та рухом від добровільних принципів до обов'язкових норм, від фрагментарних ініціатив до системних регуляторних

режимів, від технологічно нейтрального регулювання до специфічних норм, адресованих саме системам ШІ;

система наукових знань про позаєвропейські регіональні ініціативи у сфері міжнародно-правового регулювання ШІ (Африканського Союзу, АСЕАН, ОАД, Ради співробітництва арабських держав Перської затоки та інших): обґрунтовано, що кожен регіон привносить у глобальний дискурс власні цінності та регуляторні підходи – від філософії *Ubuntu* в документах Африканського Союзу та ісламських етичних концепцій в ініціативах держав Перської затоки до консенсусної моделі АСЕАН, – а спільним викликом, особливо відчутним для регіонів Глобального Півдня, залишається імплементаційний дефіцит, зумовлений обмеженістю інституційних та фінансових спроможностей.

Практичне значення отриманих результатів визначається актуальністю обраної теми, новизною положень і висновків. Теоретичні положення, узагальнення, висновки, пропозиції та практичні рекомендації, сформульовані у дисертації, можуть бути використані:

у науково-дослідній роботі – для подальшого наукового розроблення проблематики, пов'язаної з формуванням міжнародно-правового регулювання ШІ, а також для розвитку суміжних напрямів – міжнародного права прав людини, права міжнародних організацій, міжнародного інформаційного права. Завдяки міждисциплінарному підходу, що інтегрує правничі, технологічні та етичні аспекти, результати дослідження можуть бути використані в науково-дослідній роботі не лише з міжнародного права, а й у суміжних галузях;

у нормотворчій та дипломатичній діяльності – у процесі формування позиції України під час переговорів щодо міжнародно-правових актів у сфері регулювання ШІ, зокрема у межах участі держави в роботі ООН та її спеціалізованих установ, Ради Європи, а також інших міжнародних організацій. Дисертація може сприяти підвищенню ефективності діяльності зовнішньополітичних органів України у контексті ратифікації та імплементації Рамкової конвенції Ради Європи про штучний інтелект та адаптації національного законодавства до вимог Регламенту

(ЄС) 2024/1689 (Акту про ШІ) в умовах євроінтеграційного курсу держави. Порівняльний аналіз регіональних підходів до регулювання ШІ може бути використаний під час розбудови двосторонніх та багатосторонніх відносин України з державами різних регіонів у сфері технологічної співпраці;

у навчальному процесі – теоретичні положення та наукові результати дисертації можуть бути корисними для підготовки, перепідготовки та підвищення кваліфікації державних службовців, фахівців юридичного профілю, спеціалістів у галузі прав людини та цифрових технологій, а також під час розроблення і викладання курсів та спецкурсів з міжнародного публічного права, права міжнародних організацій, міжнародного права прав людини, міжнародного інформаційного права, правового регулювання ШІ тощо. Вони можуть стати у пригоді під час підготовки енциклопедичних і довідникових видань, науково-практичних коментарів, підручників і навчальних посібників з міжнародного права, його теорії та практики.

Апробація матеріалів дисертації. Дисертація обговорювалась на засіданні відділу міжнародного права та права Європейського Союзу Інституту держави і права імені В. М. Корецького Національної академії наук України. Результати дисертаційного дослідження, основні положення, висновки і рекомендації дисертації висвітлювались у виступах на міжнародних і всеукраїнських науково-практичних конференціях, круглих столах, серед них: міжнародна науково-практична конференція «Інтеграція України до Європейського Союзу: виклики в умовах воєнного стану. Пам'яті професора Павла Заворотька та професора Володимира Денисова» (м. Київ, 4–6 листопада 2024 р.), міжнародна науково-практична конференція «Наука в контексті сучасних викликів та перспектив» (м. Утрехт, 21 травня 2025 р.), X Міжнародна наукова конференція «Глобальні виклики та інновації: шляхи розвитку сучасної науки» (м. Одеса, 16 січня 2026 р.), X Міжнародна науково-теоретична конференція «Сучасний стан розвитку світової науки: характеристика та особливості» (м. Лісабон, 23 січня 2026 р.).

Публікації. Основні результати і висновки дисертаційного дослідження опубліковано в 4 статтях у фахових виданнях України з юридичних наук, а також у 4 тезах наукових доповідей на міжнародних науково-практичних конференціях.

Структура та обсяг дисертації. Структура дисертації зумовлена специфікою мети, завдань та предмету дослідження. Дисертаційна робота складається з переліку умовних позначень, вступу, трьох розділів, які поділені на дев'ять підрозділів, висновків, списку використаних джерел. Загальний обсяг дослідження становить 270 сторінок, із них список використаних джерел (290 найменувань) – 43 сторінки.

РОЗДІЛ 1. ХАРАКТЕРИСТИКА ШТУЧНОГО ІНТЕЛЕКТУ ЯК ОБ'ЄКТА МІЖНАРОДНО-ПРАВОВОГО РЕГУЛЮВАННЯ

1.1. Генеза та еволюція міжнародно-правового регулювання штучного інтелекту

Формування концепції штучного інтелекту в міжнародно-правовому дискурсі є результатом тривалої еволюції, що охоплює понад сім десятиліть технологічного розвитку, наукових досліджень та осмислення необхідності правового регулювання. Для з'ясування сучасного стану міжнародно-правового регулювання ШІ необхідно простежити історичну траєкторію розвитку самої концепції штучного інтелекту – від суто наукової ідеї до об'єкта міжнародно-правового регулювання.

Точкою відліку в історії штучного інтелекту можна вважати літню конференцію в Дартмутському коледжі 1956 року, організовану Дж. Маккарті, М. Мінським, Н. Рочестером та К. Шенноном. Саме на цій конференції Дж. Маккарті вперше запропонував термін «*artificial intelligence*» (штучний інтелект), визначивши його як науку та інженерію створення інтелектуальних машин. У пропозиції до конференції організатори висловили амбітну гіпотезу: «Дослідження ґрунтується на припущенні, що кожен аспект навчання або будь-яка інша ознака інтелекту може в принципі бути настільки точно описана, що машина зможе її симулювати» [1, с. 12]. Ця теза, на перший погляд суто технічна, закладає два концептуальні елементи, що виявилися важливими для формування сучасного міжнародно-правового регулювання ШІ. З одного боку, ідея симуляції когнітивних процесів людини порушує питання про межі автономності систем ШІ та необхідність збереження контролю людини – принцип, який згодом було закріплено як у Рекомендації ЮНЕСКО щодо етики ШІ [13], так і в Рамковій конвенції Ради Європи про штучний інтелект та права людини, демократію та верховенство права [24]. З іншого боку, визнання потенційної здатності ШІ відтворювати «будь-яку ознаку інтелекту» імпліцитно вказує на невизначеність меж розвитку таких систем, що обґрунтовує

ризикоорієнтований підхід до регулювання, запроваджений у Регламенті (ЄС) 2024/1689 (Акт про ШІ) [25]. Зазначена гіпотеза, сформульована задовго до появи перших міжнародно-правових актів у цій сфері, окреслила фундаментальну дилему, навколо якої й сьогодні будується міжнародно-правовий дискурс: технологічний потенціал ШІ є теоретично необмеженим, а отже, правове регулювання повинно передбачати можливість адаптації до непрогнозованого розвитку цієї технології.

Протягом наступних десятиліть концепція ШІ розвивалася переважно в науково-технічній площині, залишаючись поза сферою правового регулювання. Циклічність технологічного розвитку – періоди «зим штучного інтелекту» (1974–1980 та 1987–1993 роки), за яких фінансування досліджень скорочувалося через невиправдані очікування у розвитку ШІ, та періоди відродження інтересу, пов'язані з технологічними проривами – унеможлиблювала формування стабільного правового регулювання. Перші спроби осмислення напрямків правового регулювання феномену ШІ з'явилися в 1980-х роках у контексті дискусій про правовий статус експертних систем та відповідальність за рішення, прийняті з їх допомогою. Ці експертні системи були програмами, що імітували процес прийняття рішень фахівцем у конкретній вузькій галузі (медичній діагностиці, юридичному консультуванні, фінансовому аналізі тощо) на основі формалізованої бази знань та набору логічних правил типу «якщо – тоді». Саме їхнє впровадження в практику вперше поставило перед юридичною доктриною питання, що й нині залишаються центральними для міжнародно-правового регулювання ШІ: хто несе відповідальність за хибне рішення, згенероване програмою, – розробник, оператор чи користувач, і якою мірою автоматизоване рішення може замінити людське судження. Однак ці дискусії мали переважно теоретичний характер і не призвели до формування конкретних правових норм, оскільки обмежені можливості тогочасних експертних систем не створювали масштабних ризиків, які б вимагали невідкладного правового регулювання.

У 1990-х та 2000-х роках відбулася низка подій, що поступово сформували суспільний запит на правове осмислення технологій ШІ. Перемога шахового

комп'ютера Deep Blue над чемпіоном світу Г. Каспаровим у 1997 році стала першою подією у сфері ШІ, що привернула широку суспільну увагу та продемонструвала здатність системи ШІ перевершити людину в інтелектуальній діяльності, що традиційно вважалася виключно людською прерогативою. Наступне десятиліття ознаменувалося масовим впровадженням ШІ у повсякденне життя через пошукові системи, рекомендаційні алгоритми соціальних мереж та системи автоматизованої торгівлі на фінансових ринках, що поставило перед юридичною наукою питання захисту персональних даних, алгоритмічної прозорості та відповідальності за автоматизовані рішення. У 2011 році система IBM Watson перемогла чемпіонів телевікторини Jeopardy!, продемонструвавши здатність до обробки природної мови та оперування неструктурованими знаннями, що актуалізувало дискусію про застосування таких систем у різних сферах суспільної діяльності. Нарешті, перемога системи AlphaGo від DeepMind над чемпіоном світу з го Лі Седолем у 2016 році засвідчила якісно новий рівень можливостей ШІ: на відміну від шахів, го вимагає інтуїтивної оцінки позиції, що до того вважалось виключно когнітивною здатністю людини.

Сукупність цих технологічних зрушень, що перетворили ШІ із лабораторної технології на чинник, здатний суттєво впливати на суспільні відносини, у поєднанні з соціально-політичними та геополітичними факторами зумовила початок систематичних спроб формування міжнародно-правового регулювання ШІ у середині 2010-х років. На соціально-політичному рівні усвідомлення ризиків використання ШІ, порушення приватності та маніпулювання суспільною думкою стимулювало громадянське суспільство та правозахисні організації вимагати регуляторних рішень. На геополітичному рівні ШІ дедалі більше сприймається державами як технологія стратегічного значення, що впливає на економічну конкурентоспроможність та обороноздатність держав, викликавши потребу у міжнародно-правовому регулюванні його розробки та застосування для запобігання виникненню міждержавних конфліктів. Проникнення систем ШІ, здатних функціонувати з різним ступенем самостійності без безперервного втручання

людини, у критичні сфери суспільного життя актуалізувало питання юридичної відповідальності та захисту прав людини. Спроби концептуалізувати вплив ІІІ на міжнародне право знайшли відображення, зокрема, в дослідженні М.М. Мааса, який виділяє три типи такого впливу: *legal development* (розвиток нових норм для регулювання ІІІ), *legal displacement* (заміщення існуючих правил через автоматизацію процесів) та *legal destruction* (руйнування традиційних правових інститутів через фундаментальну зміну умов їхнього функціонування) [2, с. 38]. Ця триступенева типологія дозволяє систематизувати різноманітні процеси впливу технологій ІІІ на міжнародне право – від стимулювання нормотворчості до необхідності переосмислення адекватності існуючих правових інститутів новим технологічним реаліям. Глибину цих процесів підкреслює українська дослідниця Н.О. Попова, зазначаючи, що «можливості, стрімкий розвиток технологій та широкий спектр застосування ІІІ породили якісно новий вид суспільних відносин» [3, с. 245], що, відповідно, вимагає формування системного міжнародно-правового регулювання.

Зазначені фактори актуалізували фундаментальну проблему, що супроводжує правове регулювання будь-яких нових технологій, – відставання правового регулювання від технологічного розвитку. Ця проблема має глибоке теоретичне обґрунтування в науковій літературі. Британський дослідник Д. Колінґрідж ще у 1980 році у праці «Соціальний контроль технологій» сформулював парадокс, що дістав назву «дилема контролю»: «Спроби контролювати технологію, як правило, мають один із двох результатів. Або вони відбуваються занадто пізно, коли технологія вже настільки усталена, що контроль є складним, або вони відбуваються занадто рано, коли технологія ще недостатньо зрозуміла, а наслідки важко передбачити» [78, с. 11]. Сутність цієї дилеми полягає у структурній суперечності між двома умовами ефективного регулювання: на ранній стадії розвитку технології її соціальні наслідки не можуть бути достовірно спрогнозовані, що унеможливорює обґрунтоване регулювання; натомість коли негативні наслідки

стають очевидними, технологія вже глибоко інтегрована в економічну та соціальну інфраструктуру, і її контроль стає надзвичайно складним та витратним.

Сформульована Колінгріджем дилема є актуальною і для міжнародно-правового регулювання штучного інтелекту. Як зазначає М. Шерер, аналізуючи проблемні характеристики ШІ з регуляторної перспективи, розробка систем ШІ може бути «прихованою (потребувати мінімальної фізичної інфраструктури), дискретною (різні компоненти системи ШІ можуть проєктуватися без свідомої координації), дифузною (десятки осіб у географічно розпоршених місцях можуть брати участь у проєкті ШІ) та непрозорою (зовнішні спостерігачі можуть бути не в змозі виявити потенційно шкідливі характеристики системи ШІ)» [31, с. 356-357]. Ці особливості суттєво ускладнюють як превентивне (*ex ante*), так і ретроспективне (*ex post*) правове регулювання порівняно з іншими технологіями, що створювали суспільні ризики в минулому. На відміну від ядерних технологій чи фармацевтичної продукції, де об'єкт регулювання є фізично локалізованим та піддається інспекційному контролю, системи ШІ функціонують у розподіленому цифровому середовищі, а їхня поведінка може змінюватися в процесі самонавчання вже після розгортання.

Додатковим виміром цієї проблеми є часова асиметрія між логікою технологічного розвитку та логікою міжнародного нормотворення. Л. Даунс влучно узагальнив цю асиметрію: «технологія змінюється експоненціально, тоді як соціальні, економічні та правові системи змінюються інкрементально» [79, с. 8]. Якщо цикл розробки систем ШІ вимірюється місяцями (як це продемонструвало стрімке поширення генеративного ШІ у 2022–2023 роках), то класичний конвенційний процес – від початку переговорів до набрання чинності міжнародним договором – зазвичай розтягується на роки чи навіть десятиліття. Ця невідповідність ставить під сумнів придатність традиційних механізмів міжнародного нормотворення для адекватного реагування на виклики, породжувані розвитком ШІ. Як буде показано далі в цьому підрозділі та детальніше досліджено у Розділі 2, саме

ця проблема спонукала міжнародні організації до вироблення більш адаптивних моделей нормотворчої діяльності у сфері регулювання ШІ.

Усвідомлення відставання правового регулювання від технологічного розвитку зумовило закономірне домінування актів «м'якого права» на початковому етапі формування міжнародно-правового регулювання ШІ. Ця закономірність не є унікальною для сфери ШІ. Досвід правового регулювання інших технологічних сфер – біотехнологій, нанотехнологій, кіберпростору – демонструє подібну закономірність: необов'язкові акти (рекомендації, принципи, етичні настанови, керівництва), як правило, передують юридично обов'язковим нормам у регулюванні нових технологій. Крім того, необов'язкові акти створюють можливість для паралельної розробки регуляторних підходів: різні міжнародні організації можуть одночасно формулювати та апробувати різні моделі регулювання, тоді як традиційний договірний процес вимагає тривалого узгодження позицій та досягнення консенсусу, що суттєво обмежує швидкість реагування на технологічні зміни. Водночас необов'язкові акти виконують і важливу підготовчу функцію: вони дозволяють напрацювати термінологічний апарат, сформулювати консенсус навколо базових принципів та апробувати регуляторні підходи, перш ніж останні будуть закріплені у юридично обов'язкових міжнародних договорах. Цю траєкторію – від необов'язкових актів до юридично обов'язкових зобов'язань – наочно демонструє еволюція регулювання ШІ у рамках Ради Європи: від Рекомендації CM/Rec(2020)1 щодо алгоритмічних систем до Рамкової конвенції 2024 року, а також у праві ЄС: від Резолюції Європарламенту 2017 року до Регламенту (ЄС) 2024/1689, як буде детально показано у Розділі 3 цього дослідження.

Важливою передумовою формування міжнародно-правового регулювання ШІ стало прийняття національних стратегій провідними державами, що створило необхідну політичну волю для координації зусиль на міжнародному рівні. Канада у березні 2017 року стала першою державою у світі, що прийняла національну стратегію з ШІ – Панканадську стратегію штучного інтелекту, зосереджену на підтримці фундаментальних досліджень та залученні талантів. У липні того ж року

Китайська Народна Республіка оприлюднила План розвитку штучного інтелекту нового покоління, що визначив амбітну мету – досягнення світового лідерства у сфері ШІ до 2030 року. У 2018 році Франція представила національну стратегію, підготовлену на основі доповіді математика та депутата парламенту С. Вілані, а у 2019 році адміністрація Президента США видала Виконавчий указ про підтримку американського лідерства у сфері ШІ. Станом на 2026 рік національні стратегії з ШІ прийняли понад 60 держав [8], причому перша хвиля (2017–2020) охопила переважно технологічно розвинені країни, тоді як у подальшому до цього процесу долучилися й держави, що розвиваються, зокрема з Африки, Латинської Америки та Південно-Східної Азії.

Цей процес мав подвійне значення для формування міжнародно-правового регулювання. З одного боку, він продемонстрував визнання державами стратегічної важливості ШІ та необхідності його правового регулювання, створивши політичну основу для переговорного процесу. З іншого боку, різноманітність обраних національних підходів – від ціннісно-орієнтованого європейського до ринково-орієнтованого американського та державоцентричного китайського – вже на цьому етапі окреслила контури тієї фрагментації регуляторних моделей, що детально аналізується у цьому дослідженні.

На тлі цієї фрагментації національних підходів особливого значення набула діяльність міжнародних організацій та наднаціональних об'єднань, які розпочали активний процес розробки правового регулювання ШІ, спрямований на формування спільних принципів та стандартів. Важливість цього процесу підкреслює українська вчена Н.М. Оніщенко, справедливо зазначаючи, що «одним з пріоритетних напрямів розвитку нормотворчості на сучасному етапі є врахування досягнень новітніх технологій, пов'язаних зі штучним інтелектом» [4, с. 1]. Хоча детальний аналіз правової природи та змісту міжнародно-правових актів у цій сфері буде здійснений у подальших розділах дисертації, у межах цього підрозділу доцільно простежити основні етапи формування міжнародно-правового регулювання ШІ, зберігаючи послідовність від універсального рівня до регіонального, оскільки саме на

універсальному рівні було сформовано загальні принципи, що стали орієнтиром для подальших ініціатив.

Система ООН відіграла ключову роль у формуванні засад регулювання ШІ на універсальному рівні. Центральне місце в цьому процесі належить Генеральній Асамблеї ООН як головному дорадчому та представницькому органу з універсальним членством. Як зазначається у нашій попередній публікації, «міжнародно-правове регулювання штучного інтелекту на рівні ГА ООН пройшло шлях еволюції від опосередкованого фрагментарного до безпосереднього розгорнутого регулювання» [15, с. 294]: від перших опосередкованих згадок цифрових технологій у Резолюції A/RES/70/1 (2015) до прийняття 21 березня 2024 року одностайно всіма 193 державами-членами ООН Резолюції A/RES/78/265 «Використання можливостей безпечних, захищених і надійних систем штучного інтелекту для сталого розвитку» [19] – першої резолюції Генеральної Асамблеї, спеціально присвяченої штучному інтелекту. Як зазначав Генеральний секретар ООН Антоніу Гутерреш, «вперше кожна країна матиме місце за столом переговорів про ШІ» [20]. Подальшим кроком став Пакт про майбутнє (Резолюція A/RES/79/1), складовою якого є Глобальний цифровий договір (ГЦД) з окремим розділом про штучний інтелект [21]. Принциповою відмінністю ГЦД від попередніх актів ГА ООН є закріплення підходу за участю багатьох зацікавлених сторін (англ. – *multistakeholder approach*), що передбачає залучення не лише держав, а й приватного сектору, наукової спільноти та громадянського суспільства до процесу вироблення міжнародних стандартів у сфері ШІ. Детальний аналіз корпусу з 23 резолюцій ГА ООН за період 2015–2025 років, включаючи їх періодизацію, тематичну еволюцію та нормативне значення, проведено у підрозділі 2.1 цього дослідження.

Паралельно з нормотворчою діяльністю Генеральної Асамблеї розгорталася робота на рівні Секретаріату ООН та спеціалізованих установ. У 2018 році Генеральний секретар ООН Антоніу Гутерреш опублікував Стратегію щодо нових технологій, в якій ШІ визначався як «одна з ключових технологій, що матиме трансформаційний вплив на досягнення Цілей сталого розвитку» [11, с. 8],

окресливши тим самим стратегічні рамки для подальшої діяльності всієї системи ООН у цій сфері. ЮНЕСКО, як спеціалізована установа ООН з питань освіти, науки та культури, взяла на себе провідну роль у розробці етичних стандартів. Результатом безпрецедентного за масштабом процесу консультацій, що залучив понад 50 000 учасників з усього світу, стало прийняття в листопаді 2021 року Рекомендації ЮНЕСКО про етику штучного інтелекту, схваленої всіма 193 державами-членами організації [13]. Рекомендація запропонувала антропоцентричну концепцію, що інтегрує ШІ у ширший контекст культурного різноманіття, освіти та наукової етики. Як зазначає А. Каушкік, «хоча Рекомендація є добровільною та не має обов'язкової юридичної сили, вона означає *consensus ad idem* серед усіх 193 держав-членів ЮНЕСКО, що створює міцну основу для легітимності норм» [14]. Поряд з ЮНЕСКО, значну роль у формуванні міжнародних стандартів у сфері ШІ відіграють й інші спеціалізовані установи ООН в межах своїх мандатів: Міжнародний союз електров'язку (МСЕ) зосередився на технічних стандартах та координації міжнародних зусиль через платформу «AI for Good» [29]; Всесвітня організація охорони здоров'я (ВООЗ) розробила серію керівництв щодо застосування ШІ в охороні здоров'я; Всесвітня організація інтелектуальної власності (ВОІВ) ініціювала дискусію щодо впливу ШІ на систему інтелектуальної власності; Міжнародна організація праці (МОП) дослідила наслідки автоматизації для ринку праці та соціального захисту. Змістовний аналіз ініціатив спеціалізованих установ ООН здійснено у підрозділі 2.2 цього дослідження.

На міжрегіональному рівні значну роль у формуванні стандартів регулювання ШІ відіграла Організація економічного співробітництва та розвитку (далі – ОЕСР). У травні 2019 року Рада ОЕСР прийняла Рекомендацію про штучний інтелект, що визначила п'ять ключових принципів для надійного ШІ, які охоплюють орієнтацію на добробут людини та планети, повагу до верховенства права та прав людини, прозорість, надійність та безпеку функціонування, а також підзвітність організацій та осіб, що розробляють та експлуатують технології ШІ [9, с. 7–8]. Значення цієї Рекомендації полягає в тому, що вона вперше визначила основні принципи

регулювання ШІ та набула ролі концептуальної основи для подальших ініціатив на міжрегіональному та регіональному рівнях. Зокрема, принципи закладені Рекомендацією ОЕСР стали основою для Принципів ШІ G20, прийнятих на саміті в Осаці в червні 2019 року [10, с. 1], поширивши таким чином свою дію на такі держави, як Китай, Індія та Саудівська Аравія, які не є членами ОЕСР. Важливою характеристикою Рекомендації ОЕСР є передбачений нею механізм оновлення змісту відповідно до технологічного розвитку без необхідності прийняття нового акта – адаптивність, яка вже була реалізована на практиці рішенням Ради міністрів ОЕСР від 5 березня 2024 року. Детальний аналіз ролі ОЕСР та інших міжрегіональних ініціатив у формуванні міжнародних стандартів регулювання ШІ проведено у підрозділі 2.3.

Підхід за участю багатьох зацікавлених сторін, послідовно закріплений у зазначених вище актах, знайшов інституційне втілення у створенні спеціалізованих платформ для координації регуляторних зусиль. Глобальне партнерство зі штучного інтелекту (GPAI), засноване у червні 2020 року з ініціативи Франції та Канади, є міжрегіональним форумом, що об'єднує 29 держав-членів та Європейський Союз для обміну практиками та координації регуляторних підходів у сфері відповідального використання ШІ [28]. Крім того, онлайн-платформа «AI for Good» Міжнародного союзу електрозв'язку, що об'єднує понад 40 установ системи ООН, забезпечує координацію спеціалізованих установ навколо спільної мети – використання технологій ШІ для досягнення Цілей сталого розвитку [29]. Ці ініціативи демонструють еволюцію міжнародно-правового регулювання ШІ від традиційних міждержавних механізмів до багаторівневої моделі, що залучає різноманітних учасників: міжнародні організації, держави, приватний сектор та наукову спільноту.

На регіональному рівні формування правової бази для регулювання ШІ розпочалося з ініціатив Ради Європи та Європейського Союзу, які продемонстрували два відмінні, але взаємодоповнюючі підходи. Рада Європи – організація, що об'єднує 46 європейських держав навколо цінностей прав людини,

демократії та верховенства права – зосередилася на захисті основоположних прав у контексті розвитку ШІ. 8 квітня 2020 року Комітет Міністрів Ради Європи прийняв Рекомендацію CM/Rec(2020)1 державам-членам щодо впливу алгоритмічних систем на права людини [23], яка фактично охоплює широкий спектр технологій ШІ. Застосування поняття «алгоритмічні системи» замість «штучний інтелект» відображає ширший підхід Ради Європи до предмета регулювання, що охоплює як системи ШІ, так і інші форми автоматизованого прийняття рішень. Ключовою особливістю підходу Ради Європи до регулювання ШІ є акцент на потенційних загрозах для прав людини, закріплених у Європейській конвенції з прав людини: право на справедливий суд, приватність та захист даних, свободу вираження поглядів, право на рівне ставлення [23, с. 4].

Логічним розвитком цього підходу стало прийняття 17 травня 2024 року Рамкової конвенції про штучний інтелект та права людини, демократію та верховенство права (CETS No. 225), що була відкрита для підписання 5 вересня 2024 року і стала першим юридично обов'язковим міжнародним договором про штучний інтелект [24]. Конвенція встановлює систему зобов'язань держав-учасниць, що охоплює повагу до прав людини (ст. 3), дотримання демократичних цінностей (ст. 4) та верховенства права (ст. 5), а також передбачає окремі процедурні гарантії захисту осіб, які зазнають впливу технологій ШІ (ст. 10–11). Стаття 8 Конвенції встановлює обов'язок держав-учасниць «підтримувати заходи для забезпечення адекватних вимог щодо прозорості та нагляду, адаптованих до конкретних контекстів та ризиків» [24]. Важливою інновацією є положення про міжнародне співробітництво (ст. 19–23), що передбачає створення Конференції Сторін як наглядового органу. Конвенція відкрита для підписання не лише державами-членами Ради Європи, але й іншими державами, що надає їй потенційно універсального характеру. Станом на 2025 рік Конвенцію підписали 17 держав-членів Ради Європи та 10 держав, що не є її членами (зокрема, США, Канада, Японія, Ізраїль, Австралія та ін.). Україна бере активну участь у діяльності Ради Європи щодо регулювання ШІ і 5 вересня 2024 року підписала Рамкову конвенцію,

підтвердивши свій курс на інтеграцію до європейського правового простору у цій сфері. Детальний аналіз Рамкової конвенції здійснено у підрозділі 3.2.

У праві ЄС також відбувся перехід від рекомендаційних актів до юридично обов'язкового регулювання ШІ. Комітет з правових питань Європейського парламенту у 2015 році виступив із законодавчою ініціативою, результатом якої стало прийняття 16 лютого 2017 року Резолюції Європейського парламенту з рекомендаціями Комісії щодо правил цивільного права з робототехніки (2015/2103(INL)) [6]. Ця Резолюція, що має рекомендаційний характер у системі права ЄС, визначила етичні принципи розробки робототехніки та ШІ, включаючи принципи благодійності, непричинення шкоди, автономії та справедливості [6, с. 6]. Вона запровадила загальноєвропейське визначення «розумних автономних роботів», тобто роботизованих систем з елементами ШІ, здатних до самонавчання та адаптації поведінки [6, с. 5]. Ці принципи знайшли подальший розвиток у сучасному законодавстві ЄС: зокрема, принцип непричинення шкоди трансформувався в ризикоорієнтований підхід Регламенту (ЄС) 2024/1689, а вимога прозорості стала однією з ключових засад регулювання ШІ на всіх рівнях. Водночас Резолюція містила й дискусійні положення, найбільш контроверсійним з яких стала пропозиція пункту 59-f про можливість надання статусу «електронної особистості» найскладнішим автономним роботам [6, с. 12], яка, попри обґрунтовану критику наукової спільноти, каталізувала міжнародний дискурс щодо правосуб'єктності систем ШІ. Значення Резолюції полягає не у безпосередньому правовому ефекті, а у визначенні регуляторного порядку денного та формуванні концептуального апарату для подальшого законодавства ЄС.

Перехід до юридично обов'язкового регулювання ШІ у праві ЄС було здійснено прийняттям у 2024 році Регламенту (ЄС) 2024/1689 (Акт про ШІ) [25] – першого у світі комплексного горизонтального нормативно-правового акта для регулювання ШІ, що має юридично обов'язкову силу та пряму дію на території всіх держав-членів ЄС. Остаточне голосування в Європейському парламенті 13 березня 2024 року (523 голоси «за», 46 «проти», 49 утримались) підтвердило широку

підтримку нового регулювання [27]. Регламент запровадив ризикоорієнтований підхід до класифікації технологій ШІ: від систем з неприйнятним ризиком, використання яких забороняється (зокрема, системи соціального скорінгу та маніпулятивні технології), до систем з мінімальним ризиком, що не потребують додаткового регулювання. Для систем високого ризику, що застосовуються у критичних сферах – охороні здоров'я, правосудді, правоохоронній діяльності, – передбачено обов'язкову оцінку відповідності, вимоги щодо прозорості та людського нагляду [25]. Детальний аналіз Акту про ШІ здійснено у підрозділі 3.1 цього дослідження.

Рамкова конвенція Ради Європи та Регламент (ЄС) 2024/1689 (Акт про ШІ) демонструють перехід від рекомендаційних актів до юридично обов'язкових зобов'язань у регулюванні ШІ, пропонуючи при цьому дві регуляторні моделі, що склалися в рамках міжнародно-правового регулювання ШІ в цілому: модель, орієнтовану на захист прав людини та запобігання їх порушенням, і модель, засновану на класифікації ризиків, яку створює конкретна технологія ШІ. Ці моделі не є взаємовиключними, а доповнюють одна одну, що підтверджується практикою їх паралельного застосування в рамках європейського правового простору.

Розвиток міжнародно-правового регулювання штучного інтелекту демонструє складну еволюцію від підходу до нього як до суто технічного феномену до визнання його повноцінним об'єктом міжнародно-правового регулювання. Цей процес відображає не лише технологічний прогрес, але й еволюцію міжнародно-правової думки, яка намагається адаптуватися до викликів цифрової епохи, зберігаючи при цьому фундаментальні цінності міжнародного права – повагу до людської гідності, захист прав людини та сприяння міжнародному миру та безпеці.

1.2. Поняття та основні властивості штучного інтелекту як об'єкта міжнародно-правового регулювання

Формування ефективного міжнародно-правового регулювання штучного інтелекту неможливе без чіткої концептуалізації самого об'єкта регулювання. Українські дослідники Г. Машлій, О. Мосій та М. Пельчер, розглядаючи визначення терміну «штучний інтелект», обґрунтовано констатують, що «на сьогодні не існує єдиного і точного визначення досліджуваного поняття» [30, с. 80]. Проблему відсутності загальноприйнятого визначення ШІ підкреслює і професор М. Шерер, акцентуючи на тому, що це створює серйозні виклики для розробки ефективного міжнародно-правового регулювання: «Будь-який режим регулювання ШІ повинен визначати, що саме регулює цей режим; іншими словами, він повинен дати визначення штучному інтелекту» [31, с. 359].

Відсутність єдиного узгодженого визначення ШІ посилюється тим, що різні акти міжнародного права та наукові джерела послуговуються відмінним понятійним апаратом для позначення технологій, які за своєю суттю належать до сфери штучного інтелекту. Як справедливо зазначає українська дослідниця С. Корнеєва, етимологічний підхід до правового регулювання викликає труднощі «з огляду на відсутність чіткого визначення поняття ШІ, так як деякі вчені називають його «робототехніка», «машинне мислення» та таке інше» [32, с. 54]. Множинність визначень, запропонованих різними міжнародними організаціями свідчить не стільки про відсутність спроб сформулювати визначення, скільки про неузгодженість підходів, за якої кожен міжнародно-правовий акт формулює власне визначення відповідно до специфіки свого предмета регулювання. Показовим прикладом такої термінологічної різноманітності є Рекомендація CM/Rec(2020)1, яка взагалі не послуговується терміном «штучний інтелект», натомість оперуючи поняттям «алгоритмічні системи» та визначаючи їх як «застосунки, які, часто використовуючи методи математичної оптимізації, виконують одне або більше завдань, таких як збір, комбінування, очищення, сортування, класифікація та

виведення даних, а також вибір, визначення пріоритетності, створення рекомендацій та прийняття рішень» [23]. Цей приклад є показовим саме тому, що ілюструє одну з ключових проблем: міжнародно-правові акти, які за своїм предметом охоплюють технології ШІ, можуть використовувати для їх позначення цілком відмінні терміни, що ускладнює визначення меж та обсягу регулювання. За змістом наведене визначення алгоритмічних систем охоплює лише частину функцій, притаманних системам ШІ, – переважно обробку та класифікацію даних, – залишаючи поза увагою такі ключові характеристики, як автономність, здатність до навчання та генерування нового змісту. Це підтверджує, що навіть у рамках однієї міжнародної організації – Ради Європи – підходи до визначення відповідних технологій суттєво еволюціонували: від поняття «алгоритмічних систем» у Рекомендації CM/Rec(2020)1 до значно ширшого визначення «системи ШІ» у Рамковій конвенції 2024 року [24]. Саме ця множинність підходів до визначення ШІ у різних міжнародно-правових актах обумовлює необхідність їх систематичного порівняльного аналізу, який здійснюється нижче.

Складність юридичної дефініції ШІ полягає не лише в технічній багатоаспектності цього явища, але й у відсутності критеріїв, за допомогою яких можна було б визначити поняття «інтелект» поза контекстом людських когнітивних здібностей. Як влучно підкреслює Х. Шейх, «штучний інтелект настільки важко чітко визначити саме тому, що це імітація або симуляція чогось, що ми самі ще повністю не розуміємо: людського інтелекту» [45, с. 16]. Дж. Маккарті, що вперше запропонував термін «штучний інтелект», зазначав неможливість формулювання такого визначення, оскільки «ми ще не можемо загалом охарактеризувати, які саме обчислювальні процедури ми хочемо назвати інтелектуальними» [33, с. 3]. Схожої думки дотримується і М. Шерер, зазначаючи, що складність визначення ШІ полягає не в самому понятті штучності, а радше в «концептуальній неоднозначності інтелекту» [31, с. 359], адже визначення інтелекту, як правило, пов'язують з людськими характеристиками, як «єдиних істот, які загально визнано мають інтелект» [31, с. 359]. Ця невизначеність відображається і в міжнародно-правових

актах: незважаючи на численні спроби регулювання штучного інтелекту, жоден міжнародно-правовий акт не надає дефініції базового поняття «інтелект».

За таких обставин, для формування обґрунтованого визначення штучного інтелекту необхідно звернутися до міжгалузевого підходу та спершу з'ясувати сутність базового поняття «інтелект» на основі наукових досліджень у суміжних галузях знань. Оскільки вичерпний огляд усіх існуючих концепцій інтелекту виходить за межі предмета цього дослідження, для цілей формулювання міжнародно-правового визначення ШІ доцільно виокремити ті ключові ознаки інтелекту, щодо яких у сучасній науці склався відносний консенсус. На загальному рівні інтелект визначається як «здатність вчитися на досвіді та адаптуватися до нових ситуацій, розуміти та оперувати абстрактними концепціями, використовувати знання для керування власним середовищем» [34]. Дослідник диференціальної психології Р. Колом конкретизує це визначення, підкреслюючи, що «інтелект можна визначити як загальну розумову здатність до міркування, вирішення проблем та навчання, яка за своєю загальною природою інтегрує когнітивні функції, такі як сприйняття, увага, пам'ять, мова або планування» [35, с. 489]. У звіті Американської психологічної асоціації, який узагальнив результати багаторічної дискусії щодо природи інтелекту, наголошується, що «індивіди відрізняються один від одного своєю здатністю розуміти складні ідеї, ефективно адаптуватися до середовища, вчитися на досвіді, займатися різними формами міркування, долати перешкоди шляхом мислення» [36]. Незважаючи на відмінності у формулюваннях, спільним знаменником наведених визначень є посилення на когнітивні функції – «сукупність розумових процесів, пов'язаних з вирішенням проблем, плануванням, увагою, мовою та пам'яттю» [38]. У спеціалізованій літературі когнітивні функції класифікуються, як зазначає Ф.Д. Гарві, «за загальним процесом, що в них залучений, таким як пам'ять або увага, мова або виконавче функціонування» [39, с. 227], а Г. Фішер деталізує їх через п'ять первинних вимірів: увагу та виконавчі функції, функції пам'яті, мову, візуальне сприйняття та обробку, а також соціальну та емоційну обробку [40, с. 12]. Ця класифікація є релевантною для міжнародно-

правового визначення ІІІ, оскільки саме здатність до виконання окремих когнітивних функцій – обробки інформації, розпізнавання закономірностей, прийняття рішень, генерування мовних конструкцій – є тим критерієм, за яким системи ІІІ відрізняються від традиційного програмного забезпечення.

На переважаючій значимості когнітивних функцій у визначенні інтелекту наголошує також О. Баранов, який здійснив найбільш систематичну в українській правовій науці спробу адаптувати поняття когнітивних функцій до потреб правового регулювання ІІІ. Проаналізувавши множину запропонованих визначень, дослідник доходить висновку, що «швидко стає очевидною велика подібність між багатьма визначеннями» [41, с. 36]. На основі цього аналізу він встановив, що найбільш вживаними поняттями при формулюванні дефініцій інтелекту є ті, що стосуються когнітивних функцій та здібностей (когнітивні функції, когнітивні здібності, когнітивні завдання), розумових процесів (мислити, розуміти, обробка інформації, пам'ять), а також цілеспрямованої діяльності (здатність приймати рішення, навчатися, діяти ефективно) [41, с. 36]. Ключовим висновком є те, що «оскільки прийняття рішень, мислення, пам'ять, розуміння, обробка інформації, навчання є результатом реалізації когнітивних функцій, то фактично для всіх визначень терміну «інтелект» найбільш характерним є використання терміну «когнітивна функція» [41, с. 36].

Розвиваючи цей підхід, дослідник пропонує поділ когнітивних функцій на два типи, що має безпосереднє значення для диференціації систем ІІІ за рівнем складності. По-перше, проста когнітивна функція визначається як «специфічна діяльність мозку людини із здійснення окремих простих операцій щодо збору, створення, перетворення, обробки та накопичення (збереження) інформації» [41, с. 39]. По-друге, складна когнітивна функція розуміється як «певна сукупність простих когнітивних функцій мозку людини призначена для виконання автономної складної операції із інформацією, яка має самостійне значення» [41, с. 39]. Цей поділ є релевантним для міжнародно-правового регулювання, оскільки дозволяє чіткіше окреслити, які саме когнітивні функції відтворюються сучасними системами ІІІ:

системи вузького ШІ, як правило, виконують окремі прості когнітивні функції або їх обмежені комбінації, тоді як перспективні системи загального ШІ мали б інтегрувати складні когнітивні функції, що має наслідки для визначення обсягу та інтенсивності регулювання.

Виявлене значення когнітивних функцій для визначення інтелекту дозволяє перейти до формулювання поняття «штучний інтелект». О. Баранов визначає інтелект як «повну сукупність складних та простих когнітивних функцій мозку, призначених для визначення найкращого варіанту рішення щодо цілеспрямованої та ефективної поведінки чи соціальної активності у будь-яких конкретних умовах та середовищах» [41, с. 39]. Принциповим для концепції штучного інтелекту є висновок дослідника про те, що різні люди здатні демонструвати однакові результати когнітивної діяльності, незважаючи на індивідуальні відмінності у будові мозку. Це означає, що для відтворення інтелектуальних функцій вирішальне значення має не біологічна природа носія, а здатність досягати еквівалентних результатів – тобто виконувати ті самі когнітивні функції іншими засобами. Саме цей принцип лежить в основі самої ідеї штучного інтелекту: якщо когнітивні функції можуть бути відтворені небіологічною системою з еквівалентним результатом, така система може кваліфікуватися як штучний інтелект.

Для формулювання міжнародно-правового визначення ШІ необхідно також з'ясувати зміст другого елемента цього терміна – поняття «штучний». Це питання має практичне значення, оскільки від розуміння критеріїв «штучності» залежить, які саме системи підпадуть під відповідний правовий режим. О. Баранов, аналізуючи використання поняття «штучний» у різних контекстах (штучні магніти, серце, волокна, супутник Землі тощо), доходить висновку, що визначальним критерієм є функціональне призначення об'єкта: «при вирішенні проблеми вибору еквівалентних характерних властивостей певних природних/штучних об'єктів необхідно в першу чергу враховувати їх функціональне призначення або звертати увагу на їх функціональність в процесі досягнення певної мети» [41, с. 35]. На цій основі дослідник формулює визначення: «штучний – це речовина, матеріал, об'єкт,

процес чи середовище, що створено як копія, модель, імітація або аналогія природного з, як мінімум, еквівалентними характерними властивостями, наявність яких обумовлює певну користувачьку цінність» [41, с. 36].

Функціональний підхід до визначення «штучного», запропонований О. Барановим, підтверджується дослідженнями у суміжних галузях знань. Як зазначає С. Гартманн, «симуляція імітує один процес іншим процесом» [42, с. 5], а отже, штучний об'єкт є функціональним тією мірою, якою він здатний відтворювати результати, еквівалентні природному аналогу [43]. Водночас Ф. Б'янкіні звертає увагу на те, що «межа між природним і штучним не є такою чіткою та однозначною, як це було раніше» [44, с. 401], що є особливо актуальним і для систем ШІ, які здатні досягати результатів, що за своїми характеристиками дедалі менше відрізняються від результатів людської когнітивної діяльності. Для цілей міжнародно-правового регулювання ця обставина підкреслює доцільність визначення ШІ через функціональну еквівалентність результатів окремих когнітивних функцій, а не через природу носія. Водночас така еквівалентність стосується лише відтворення окремих когнітивних функцій і не означає тотожності ШІ людському інтелекту, від якого він залишається якісно відмінним.

Таким чином, аналіз базових понять «інтелект» та «штучний» дозволяє сформулювати вихідне припущення: штучний інтелект може бути визначений як система, здатна відтворювати окремі когнітивні функції людини з результатами, еквівалентними результатам реалізації людиною окремих когнітивних функцій, незалежно від природи носія. Водночас це припущення потребує співставлення з міжнародно-правовими актами, які формулюють власні визначення ШІ відповідно до специфіки предмета свого регулювання. Різні міжнародні організації та міждержавні об'єднання, зокрема ООН, ОЕСР, ЮНЕСКО, Рада Європи, Європейський Союз та інші – розробили власні визначення штучного інтелекту, кожне з яких орієнтується на специфіку предмета регулювання відповідного акта. У Глобальному цифровому договорі держави-члени ООН визнали, що «міжнародне регулювання штучного інтелекту вимагає гнучкого, міждисциплінарного та

адаптивного підходу за участю багатьох зацікавлених сторін» [21, п. 54], що свідчить про усвідомлення відсутності єдиного концептуального базису для визначення ШІ. Саме потреба у різнобічному вивченні ШІ з метою вироблення єдиного й повного його визначення обумовлює доцільність системного порівняльного аналізу наявних підходів, який здійснюється далі.

За таких умов визначення ШІ, сформульовані в різних актах стають теоретичною основою для систематизації існуючих підходів та виявлення підґрунтя для вироблення єдиного підходу до визначення ШІ та режиму його міжнародно-правового регулювання. На основі аналізу міжнародно-правових актів ООН, ОЕСР, ЮНЕСКО, Ради Європи, Європейського Союзу та інших міжнародних організацій, а також наукової літератури з досліджуваного питання, пропонуємо власну класифікацію підходів до визначення ШІ. Критеріями виокремлення підходів є ті ознаки ШІ, що характеризують його як особливий феномен і покладені в основу відповідного визначення: технічна природа системи, її вплив на людину, виконувані функції, системна організація, динамічний характер, цілеспрямованість діяльності або сукупність цих ознак. Відповідно, виокремлено сім підходів: технологічний, що розглядає ШІ, передусім, через призму технічних методів та алгоритмів; антропоцентричний, що ставить у центр визначення вплив ШІ на людину та суспільство; функціональний, що концентрується на операціях, які виконують системи ШІ; системний, що розглядає ШІ як складну соціотехнічну систему; еволюційний, що підкреслює динамічний характер розвитку цієї технології; цілеорієнтований, що базується на концепції досягнення цілей як центрального елемента поведінки ШІ; та інтегративний, що синтезує елементи зазначених підходів. Далі кожен з цих підходів розглядається на основі аналізу конкретних міжнародно-правових актів, що їх втілюють.

Технологічний підхід розглядає ШІ передусім через призму технічних методів та алгоритмів, покладених в основу функціонування системи. Елементи цього підходу простежуються, зокрема, у визначенні, запропонованому Експертною групою високого рівня з питань штучного інтелекту при Європейській комісії (AI

HLEG) у 2019 році, яке описує ШІ як «програмні (а можливо, також апаратні) системи, розроблені людьми, які за наявності складної мети діють у фізичному або цифровому вимірі, сприймаючи своє середовище шляхом збору даних, інтерпретуючи зібрані структуровані або неструктуровані дані, міркуючи на основі знань або обробляючи інформацію, отриману з цих даних, та вирішуючи найкращі дії для досягнення заданої мети» [47, с. 36]. Хоча це визначення містить і функціональні елементи (сприйняття середовища, міркування), його ядро становить саме опис технічних процесів обробки даних. Водночас слід зазначити, що AI HLEG підготувала це визначення як окремий супровідний документ до Етичних настанов для надійного ШІ [47], які, у свою чергу, закладають людиноцентричний підхід до регулювання, що буде розглянуто нижче. Ця обставина ілюструє, що технологічний та антропоцентричний підходи можуть поєднуватися навіть у межах однієї інституції, взаємодоповнюючи один одного та виконуючи різні функції: перший слугує ідентифікації об'єкта регулювання, другий – визначенню засад його регулювання. Ознаки технологічного підходу містить також Рекомендація CM/Rec(2020)1, що оперує поняттям «алгоритмічні системи» та визначає їх через перелік конкретних технічних операцій – збір, сортування, класифікацію та виведення даних [23]. Показово, що серед сучасних юридично обов'язкових міжнародно-правових актів суто технологічний підхід не знайшов нормативного закріплення: і Регламент (ЄС) 2024/1689 (Акт про ШІ), і Рамкова конвенція Ради Європи про ШІ обрали ширші визначення, що свідчить про усвідомлення обмежень технологічного підходу на міжнародному рівні.

З позицій науки міжнародного права технологічний підхід має суттєві обмеження. Надмірна прив'язка визначення до конкретних технічних методів створює ризик швидкого застарівання правових норм, оскільки технології ШІ розвиваються значно швидше, ніж відбувається перегляд міжнародно-правових актів. Крім того, такий підхід залишає поза увагою вплив ШІ на суспільні відносини та основоположні права людини, що є критично важливим для міжнародно-правового регулювання. Як наголошується в Рекомендації ЮНЕСКО щодо етики

штучного інтелекту, регулювання має базуватися на «людській гідності та правах людини» [13], що виходить за межі суто технологічного розуміння феномену ШІ. Саме усвідомлення цих обмежень зумовило формування альтернативних підходів до визначення ШІ.

На противагу технологічному, антропоцентричний підхід визначає штучний інтелект через призму його впливу на людину, суспільство та міжнародний правопорядок. Цей підхід відображає фундаментальну орієнтацію міжнародного права на захист людської гідності як однієї з основоположних цінностей сучасного міжнародного права. Антропоцентричний підхід виходить з того, що юридична дефініція ШІ має базуватися передусім на наслідках застосування цих технологій для реалізації прав людини, соціальної справедливості та верховенства права, а не на технічних характеристиках систем.

Нормативне втілення антропоцентричного підходу знайшло відображення в низці актів міжнародних організацій. Зокрема, Рекомендація ЮНЕСКО щодо етики штучного інтелекту наголошує на свідомому відході від спроб створення єдиної технічної дефініції, обґрунтовуючи це динамічним характером технологічного розвитку: «Рекомендація не має на меті надати єдине визначення ШІ, оскільки таке визначення з часом мало б змінюватися відповідно до технологічного розвитку. Швидше, її метою є вирішення тих особливостей систем ШІ, які мають центральне етичне значення» [13, с. 3]. Натомість Рекомендація зосереджується на етично значущих характеристиках систем ШІ через концепцію «життєвого циклу» технології, визначаючи системи ШІ як «технології обробки інформації, що інтегрують моделі та алгоритми, які створюють здатність до навчання та виконання когнітивних завдань, що призводять до таких результатів, як прогнозування та прийняття рішень у матеріальному та віртуальному середовищах ... Етичні питання щодо систем ШІ стосуються всіх етапів життєвого циклу системи ШІ, який тут розуміється як такий, що охоплює дослідження, проектування та розробку до розгортання та використання, включаючи технічне обслуговування,

експлуатацію, торгівлю, фінансування, моніторинг та оцінку, валідацію, завершення використання, розбирання та припинення використання» [13, с. 4].

Антропоцентричний підхід також знайшов своє нормативне втілення у концепції людиноцентричного ШІ (англ. – human-centric AI), що була формально сформульована тією ж групою AI HLEG, але вже не у визначенні ШІ, а в Етичних настановах для надійного ШІ. Відповідно до цих Настанов, «системи ШІ повинні бути людиноцентричними, ґрунтуючись на зобов'язанні щодо їх використання на службі людства та спільного блага, з метою покращення добробуту та свободи людини» [47, с. 4]. Концепція ґрунтується на чотирьох етичних принципах: повазі до людської автономії, запобіганні шкоді, справедливості та пояснюваності [47, с. 12], та передбачає, що системи ШІ повинні відповідати семи ключовим вимогам надійності: людський контроль та нагляд, технічна надійність та безпека, приватність та управління даними, прозорість, різноманітність та недискримінація, суспільний та екологічний добробут, підзвітність [47, с. 14]. Ця концепція стала наріжним каменем підходу ЄС до регулювання ШІ і набула подальшого розвитку в Регламенті (ЄС) 2024/1689 (Акті про ШІ), преамбула якого прямо закріплює зазначений підхід як основу регулювання [25].

Сильною стороною антропоцентричного підходу є його відповідність основоположним принципам міжнародного права, зокрема загальній повазі та захисту прав людини та та можливість впровадження цінностей, принципів і норм міжнародного права прав людини у регулювання ШІ. Цей підхід забезпечує спадкоємність правового регулювання, дозволяючи поширити вже апробовані механізми захисту прав людини на нові технологічні реалії. Водночас необхідно визнати певні методологічні обмеження цього підходу. По-перше, акцентування на впливі ШІ на права людини може призводити до розмитості правових дефініцій, оскільки поняття «вплив на права людини» є оціночним і залежить від конкретних часу, місця, обставин та сфери застосування технології. По-друге, антропоцентричний підхід може ускладнювати визначення чітких юридичних критеріїв для кваліфікації систем як ШІ. Наприклад, проста рекомендаційна система

та складна система підтримки судових рішень по-різному впливають на права людини, проте критерій «впливу на людину» сам по собі не дає змоги однозначно встановити, яка з них є системою ШІ у розумінні відповідного правового режиму. По-третє, визначення ШІ виключно через його вплив на людину та суспільство не дозволяє чітко окреслити межі поняття: будь-яка технологія тією чи іншою мірою впливає на права людини: за таким критерієм до систем ШІ довелось б віднести і звичайний пошуковий сервіс, що робить визначення надмірно широким та позбавляє його розмежувальної функції.

Функціональний підхід до визначення штучного інтелекту концентрується на виконуваних ним операціях, відокремлюючи їх від технічної архітектури ШІ та його впливу на людину і суспільство. Цей підхід має глибоке методологічне коріння в юридичній науці: функціональний аналіз традиційно використовується для визначення правового режиму різних явищ та інститутів, дозволяючи застосовувати аналогію права в ситуаціях правових прогалин.

На основі аналізу міжнародно-правових актів можна виокремити ключові операції, через які функціональний підхід характеризує системи ШІ: сприйняття та аналіз інформації з навколишнього середовища, навчання на основі досвіду, прийняття рішень для досягнення певних цілей, адаптація поведінки відповідно до змін середовища, прогнозування майбутніх подій та станів, генерування нового контенту або рішень. Регламент (ЄС) 2024/1689 (Акт про ШІ), що закріплює людиноцентричний підхід як загальну засаду регулювання, у самому визначенні системи ШІ демонструє саме елементи функціонального підходу, описуючи її як «машинну систему, яка призначена для роботи з різним рівнем автономії і яка може, для явних або неявних цілей, генерувати результати, такі як прогнози, рекомендації або рішення, що впливають на фізичне або віртуальне середовище» [25]. Таке визначення акцентує увагу саме на функціональних можливостях систем ШІ – здатності генерувати результати певного типу – а не на технічній архітектурі чи методах їх досягнення. Це підтверджує спостереження, зроблене при аналізі технологічного підходу: один і той самий міжнародно-правовий акт може

поєднувати різні підходи, використовуючи функціональне визначення для ідентифікації об'єкта регулювання та антропоцентричний підхід для визначення ціннісних засад цього регулювання. У зарубіжній правовій доктрині функціональний підхід знайшов найбільш лаконічне вираження у визначенні М. Шерера, який описує ШІ як «машини, здатні виконувати завдання, які, якби їх виконувала людина, вимагали б інтелекту» [31, с. 362]. Це визначення ілюструє сутність функціонального підходу: замість опису технічної архітектури чи впливу на суспільство воно відсилає безпосередньо до результату діяльності системи, що робить його зручним для цілей правозастосування.

Функціональний підхід до дефініції ШІ має ряд переваг. По-перше, він дозволяє створювати технологічно нейтральні норми, які мають залишатися актуальними незалежно від технологічних змін та інновацій. Така технологічна нейтральність є критично важливою в умовах швидкого розвитку технологій ШІ, при якому конкретні технічні рішення можуть застарівати протягом короткого періоду часу. По-друге, функціональний підхід полегшує визначення правового режиму застосування ШІ на практиці, оскільки дозволяє порівнювати їх з уже врегульованими видами діяльності. Якщо система ШІ виконує функції, аналогічні тим, що традиційно виконуються людиною у певній правовій сфері, до неї можуть застосовуватися відповідні правові норми за принципом функціональної еквівалентності. По-третє, функціональний підхід дозволяє уникнути спроб визначити «справжню природу» або «сутність» штучного інтелекту як філософської категорії, зосереджуючись натомість на регулюванні конкретних функцій ШІ та їх юридично значущих наслідків.

Водночас, необхідно зазначити, що функціональний підхід має певні методологічні обмеження. Основною проблемою є те, що аналогічні функції можуть виконуватися різними технологічними засобами з принципово різними правовими наслідками. Система ШІ, що виконує ті самі функції, що й людина, може створювати специфічні ризики, пов'язані з непрозорістю процесу прийняття рішень, відсутністю свідомого розуміння контексту та обмеженою здатністю до етичного судження.

Крім того, надмірна орієнтація на функціональну схожість може призводити до ігнорування якісних відмінностей між людською та машинною діяльністю, що є важливим для правового регулювання критичних сфер застосування ШІ, таких як воєнна промисловість, національна безпека, фінанси, міграційна політика, соціальний скоринг тощо.

За системного підходу ШІ розглядається як складна соціотехнічна система, що включає технологічні, соціальні, економічні, культурні та правові компоненти в їх взаємозв'язку та взаємозалежності. Цей підхід базується на методології системного аналізу, згідно з якою ШІ не існує ізольовано, а функціонує в складному контексті міжнародних відносин, впливаючи на різні рівні правового регулювання – від індивідуального до глобального.

Так, українська дослідниця Н.О. Попова пропонує системне визначення ШІ як «організованої сукупності інформаційних технологій та систем, створеної для виконання системних комплексних завдань, здатної до самостійності та самовдосконалення, яка імітує людські когнітивні здібності для аналізу отриманої інформації та створення результатів (прийняття рішень, розробка висновків, рекомендацій, прогнозів) і впливає на реальне та віртуальне середовище» [3, с. 245]. Це визначення демонструє прагнення охопити всі ключові системні характеристики феномену ШІ – організаційну структуру, функціональні можливості, автономність та соціальний вплив.

Елементи системного підходу простежуються також у Концепції розвитку штучного інтелекту в Україні (2020), яка, хоча і є національним правовим актом, є прикладом імплементації міжнародних підходів до визначення ШІ на національному рівні. Концепція визначає ШІ як «організовану сукупність інформаційних технологій, із застосуванням якої можливо виконувати складні комплексні завдання, шляхом використання системи наукових методів досліджень і алгоритмів оброблення 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» [49].

Значення системного підходу для міжнародно-правового регулювання полягає у тому, що він розглядає ШІ не як окремий технічний засіб, а як сукупність взаємопов'язаних компонентів, що охоплює організаційну структуру, функціональні можливості, автономність та вплив на середовище. Таке розуміння орієнтує нормотворця на формулювання визначень, які б охоплювали ШІ у всій його складності, а не зводили його до окремого аспекту – технічного, функціонального чи етичного.

Еволюційний підхід до визначення ШІ акцентує увагу на динамічній природі цього явища та виходить з того, що будь-яке статичне визначення ШІ приречене на швидке застарівання у зв'язку з інтенсивним технологічним розвитком. Парадоксальну природу цієї проблеми ілюструє висловлювання Дж. Маккарті: «Як тільки це починає працювати, ніхто більше не називає це ШІ» [31, с. 361]. М. Шерер ілюструє цю тезу прикладом гри в шахи: досягнення комп'ютером перемоги над найкращим гравцем у 1997 році було інтерпретовано не як доказ інтелекту машини, а як свідчення того, що ця здатність не є показником інтелекту [31, с. 361]. Критерії ШІ постійно зміщуються у міру технологічного прогресу, що створює складнощі для формування сталих правових дефініцій.

Еволюційний підхід знайшов своє відображення у концепції спектру (або континууму) штучного інтелекту, яка розглядає різні системи не як чітко розмежовані категорії, а як точки на неперервному спектрі можливостей. Концептуальне обґрунтування такого підходу знаходимо у працях дослідника у галузі штучного інтелекту Б. Гьортцеля, який зазначав, що «спектр дизайнів систем ШІ включає системи з символічними, емерджентними, гібридними та універсалістськими характеристиками» [50, с. 1]. Найбільш систематичну спробу операціоналізувати цю концепцію здійснили М. Р. Моріс, Дж. Сол-Дікштайн та Н. Фідель, запропонувавши класифікацію, що «вводить рівні продуктивності загального штучного інтелекту (ЗШІ), загальності та автономії, забезпечуючи

спільну мову для порівняння моделей, оцінки ризиків та вимірювання прогресу на шляху до ЗШІ» [51, с. 1]. Найбільш поширеною реалізацією концепції континууму є тривидова класифікація [52]: вузький штучний інтелект (ВШІ), спеціалізований у вирішенні завдань в обмеженій сфері та вже широко впроваджений у різних галузях; загальний штучний інтелект (ЗШІ), що потенційно досягає рівня людського інтелекту та характеризується здатністю до трансферу знань між різними сферами – термін, що вперше з'явився у 1997 році у роботі М. Губруда [53]; та гіпотетичний штучний суперінтелект (ШСІ), який Н. Бостром визначає як «будь-який інтелект, що значно перевищує когнітивну продуктивність людей практично в усіх сферах, що становлять інтерес» [54, с. 39].

Водночас міжнародні організації здебільшого не використовують тривидову класифікацію у нормативних актах, обираючи натомість функціональні визначення з градуюванням на основі ризиків, а не рівнів інтелекту. Це свідчить про те, що для цілей міжнародно-правового регулювання еволюційний підхід має значення не стільки як основа для класифікації типів ШІ, скільки як застереження щодо необхідності формулювати визначення достатньо гнучко, аби вони залишалися придатними в умовах технологічних змін, а також орієнтує на прогнозування подальших напрямів технологічного розвитку.

Цілеорієнтований підхід ґрунтується на концепції досягнення цілей (наприклад, перемога у грі, оптимізація маршруту чи максимізація точності прогнозу) як центрального елемента інтелектуальної поведінки. Автори класичного підручника зі штучного інтелекту С. Рассел та П. Норвіг систематизували визначення ШІ у чотири базові категорії: системи, що мислять як людина, системи, що діють як людина, системи раціонального мислення та системи раціональної дії [55, с. 2]. Водночас, центральним елементом цього підходу є конструкція «раціонального агента» – автономної системи, що здійснює дії з метою досягнення оптимального результату або, за умов невизначеності, максимально очікуваного позитивного результату [55, с. 4]. Концептуальні засади цього підходу знаходять відображення у працях Дж. Маккарті, який визначав інтелект як обчислювальний

компонент здатності досягати цілей у реальному світі [33, с. 2], а також С. Омохундро, який кваліфікує ШІ як систему, що «володіє визначеними цілями та здійснює цілеспрямовану діяльність у фізичному чи цифровому середовищі для їх досягнення» [56, с. 1].

Проте з позицій міжнародно-правового регулювання цілеорієнтований підхід стикається з концептуальними та практичними викликами. М. Шерер критично оцінює його придатність для нормотворення, зазначаючи, що він «просто замінює один термін, що важко визначити (інтелект) на інший (ціль)» [31, с. 361]. Основна проблема полягає у тому, що в юридичній мові поняття «ціль» тісно пов'язане з категорією «намір», що породжує питання щодо правосуб'єктності систем ШІ та їхньої здатності мати наміри. Крім того, визначення через поняття «ціль» створює ризик надмірно широкого тлумачення: якщо дефініція обмежується формулюванням «кінцевий результат, що може бути конкретно визначений», то навіть елементарна механічна система підпадатиме під це визначення. Водночас критерій раціональності є і надмірно широким, і недостатнім: з одного боку, комп'ютерні шахові програми «діють раціонально», проте не створюють правових ризиків; з іншого – системи ШІ, що функціонують нераціонально або непередбачувано, можуть становити серйозні загрози для прав людини, але не охоплюються цим критерієм. На нашу думку, підхід, заснований виключно на концепції цілей та раціональності дій, не забезпечує достатньої правової визначеності для міжнародно-правового регулювання ШІ.

Проведений аналіз підходів до визначення ШІ в актах міжнародного права та наукових джерелах засвідчує, що жоден із розглянутих підходів, узятий окремо, не позбавлений суттєвих обмежень. Саме виявлені обмеження кожного з них обґрунтовують доцільність інтегративного підходу, який поєднує їхні сильні сторони у комплексному правовому визначенні ШІ. Інтегративний підхід відображає усвідомлення того, що жоден окремо взятий підхід не може забезпечити вичерпного охоплення багатовимірного феномену ШІ як об'єкта міжнародно-

правового регулювання, та набуває нормативного закріплення у сучасних міжнародно-правових актах.

Найбільш послідовно інтегративний підхід втілений у Регламенті (ЄС) 2024/1689 (Акті про ШІ). Як було зазначено при аналізі функціонального підходу, визначення системи ШІ у ст. 3(1) цього Регламенту акцентує увагу на функціональних можливостях системи – здатності генерувати результати певного типу [25]. Водночас Регламент поєднує це функціональне визначення з людиноцентричним підходом, закріпленим у преамбулі, та ризикоорієнтованою класифікацією, що диференціює системи за рівнем потенційної загрози для прав людини: системи з неприйнятним ризиком (заборонені), високим ризиком (що потребують оцінки відповідності), обмеженим ризиком (з вимогами щодо прозорості) та мінімальним ризиком (без додаткових обмежень). Саме це поєднання технічних, функціональних, антропоцентричних та ризикоорієнтованих елементів у єдиній правовій конструкції відрізняє інтегративний підхід від інших, розглянутих вище.

У Пакті про майбутнє (англ. – Pact for the Future) інтегративний підхід реалізується через поєднання антропоцентричного та системного підходів, зокрема п. 52 встановлює вимогу «розвивати справедливі та інклюзивні підходи до використання переваг штучного інтелекту та зменшення ризиків при повному дотриманні норм міжнародного права, включаючи міжнародне право прав людини» [21], закріплюючи вимогу повного дотримання норм міжнародного права, у тому числі міжнародного права прав людини, при використанні ШІ. У п. 53 Пакту підкреслюється необхідність «сприяти процвітанню різноманітних культур і мов та підтримувати застосування отриманих на місцевому рівні даних на благо розвитку країн та співтовариств» [21, п. 53], що відображає соціально-правовий вимір регулювання. Системний характер інтегративного підходу закріплюється у п.54 Пакту, де визнається, що «міжнародне регулювання штучного інтелекту вимагає гнучкого, міждисциплінарного та адаптованого підходу за участю багатьох зацікавлених сторін» [21, п. 54], а у п. 55(d) Пакту наголошується на необхідності

забезпечити «транспарентність, підзвітність та ретельний нагляд за системами штучного інтелекту з боку людини» [21, п. 55(d)]. Показово, що резолюції ГА ООН послідовно застосовують конструкцію «системи штучного інтелекту», що вказує на нормативне закріплення розуміння ШІ саме як системного явища, а не окремого технологічного засобу.

Інтегративний підхід дозволяє забезпечити комплексність правового визначення ШІ, що охоплює технічні, функціональні, етичні та соціально-правові аспекти феномену штучного інтелекту, зберігаючи водночас необхідну правову визначеність через чітке окреслення нормативних критеріїв. Саме цей підхід покладено в основу авторської дефініції ШІ, що пропонуватиметься далі.

Незважаючи на множинність підходів до визначення ШІ, проведений аналіз актів міжнародного права та наукових джерел дозволяє виокремити спільні ознаки, які наскрізно присутні у різних визначеннях цього феномену, що засвідчує конвергенцію підходів щодо основних характеристик систем ШІ.

Першою такою ознакою є здатність систем ШІ до сприйняття середовища, що їх оточує. Так, у визначенні, сформульованому AI HLEG, системи ШІ описуються як такі, що «діють у фізичному або цифровому вимірі, сприймаючи своє середовище через збір даних» [47], а Рекомендація ЮНЕСКО підкреслює, що визначення ШІ «типово включає аспекти сприйняття» [13]. Другою ознакою є здатність до обробки інформації, що виходить за межі базових обчислень. Регламент (ЄС) 2024/1689 (Акт про ШІ) визначає ключову характеристику систем ШІ як їхню «здатність до інференції», що означає «процес отримання результатів, таких як прогнози, контент, рекомендації або рішення», підкреслюючи, що ця здатність «виходить за межі базової обробки даних, забезпечуючи навчання, міркування або моделювання» [25]. AI HLEG конкретизує цей процес через здатність систем «інтерпретувати зібрані структуровані чи неструктуровані дані, міркувати про знання або обробляти інформацію, отриману з цих даних» [47].

Третьою ознакою є здатність до прийняття рішень з визначеним рівнем автономності від безпосереднього людського втручання. Визначення ОЕСР,

оновлене у 2023 році порівняно з первісною редакцією Рекомендації 2019 року, характеризує системи ШІ як такі, що «виводять зі вхідного сигналу, який вони отримують, як генерувати виходи, такі як прогнози, контент, рекомендації або рішення», підкреслюючи, що «різні системи ШІ відрізняються за рівнями автономності та адаптивності після розгортання», тобто після введення системи в експлуатацію. [9]. Рекомендація ЮНЕСКО зазначає, що системи ШІ «розроблені для роботи з різними ступенями автономії за допомогою моделювання та представлення знань» [13]. Четвертою ознакою є орієнтація на досягнення конкретних цілей. Рекомендація ОЕСР та Акт про ШІ визначають системи ШІ як такі, що функціонують «для явних або неявних цілей» [9; 25], визнаючи, що цілі можуть бути як визначеними розробниками, так і виникати у процесі навчання самої системи ШІ. AI HLEG підкреслює, що системи створюються «маючи складну мету» та здійснюють дії для «досягнення заданої мети» [47].

Виявлені ознаки у визначеннях штучного інтелекту засвідчують формування спільного понятійного ядра, яке може слугувати підґрунтям для подальшого узгодження міжнародно-правового регулювання ШІ. Конвергенція підходів щодо ключових характеристик систем ШІ простежується як у рамках «м'якого права» (резолуції, рекомендації), так і у юридично обов'язкових міжнародних договорах. Водночас множинність підходів до операціоналізації цих спільних ознак у конкретних міжнародно-правових актах зумовлюється об'єктивними чинниками, зокрема різними регуляторними цілями актів, особливостями правових систем держав-учасниць, а також специфікою сфер застосування норм. Така варіативність не перешкоджає формуванню узгодженого міжнародно-правового регулювання використання ШІ, а відображає природну еволюцію правового регулювання нових технологій.

Поряд із визначенням поняття ШІ, формування ефективного міжнародно-правового регулювання потребує аналізу основних властивостей штучного інтелекту, що детермінують особливості відповідного правового режиму. Проведений аналіз концептуальних підходів до визначення ШІ, наукових джерел та

міжнародно-правових актів дозволяє виокремити ключові властивості ШІ, що обумовлюють необхідність особливого підходу до його регулювання.

Однією з ключових властивостей ШІ є транскордонність його застосування. Хоча окрема система може функціонувати й у межах однієї держави, технології ШІ за своєю архітектурою тяжіють до транскордонного використання. Як зазначається в Преамбулі Регламенту (ЄС) 2024/1689 (Акт про ШІ), «системи штучного інтелекту можуть легко розгортатися в багатьох секторах економіки та суспільства, включаючи транскордонне використання, та циркулювати по всьому Союзу» [25]. Ця властивість створює виклик для сучасного міжнародного права, побудованого на принципі територіального суверенітету держав. Транскордонність функціонування ШІ проявляється на кількох рівнях: технологічна архітектура систем ШІ часто передбачає розподілену обробку даних через хмарні обчислення, що фізично розташовані в різних юрисдикціях; алгоритми машинного навчання можуть тренуватися на даних, зібраних з різних країн, створюючи складні питання щодо застосовного права та юрисдикції; результати діяльності ШІ можуть мати екстериторіальний ефект, впливаючи на права та інтереси осіб у різних державах. Як справедливо зазначає Н.О. Попова, «використання ШІ не має державних кордонів і має більшою мірою транснаціональний характер, одним із заходів протидії зловживанням у цій сфері має стати міжнародний механізм регулювання цих суспільних відносин» [3, с. 245]. Транскордонність ШІ породжує необхідність узгодження національних підходів до його регулювання та створення міжнародних стандартів, оскільки відсутність єдиних міжнародних стандартів регулювання ШІ створює ризик «регуляторного арбітражу», за якого розробники систем ШІ можуть обирати юрисдикції з найменш обтяжливими вимогами.

Другою ключовою властивістю ШІ є його подвійне використання (англ. – dual use) – можливість застосування однієї й тієї ж технології як для цивільних, так і для воєнних цілей. Як зазначає А. Г. де Агрета, «більшість технологій штучного інтелекту мають подвійне призначення. Вони інтегровані як у мирні цивільні застосування, так і у військові системи озброєнь» [58]. Алгоритми розпізнавання

образів можуть використовуватися як для медичної діагностики, так і для систем наведення зброї; системи обробки природної мови – як для покращення комунікації, так і для створення дезінформаційного контенту. Н.О. Попова підкреслює, що «технології ШІ часто використовуються і у веденні бойових дій» [3, с. 246], і саме можливість такого застосування технологій ШІ перетворює подвійне використання з нейтральної техніко-функціональної ознаки на гостру міжнародно-правову проблему.

Найгострішого вияву ця проблема набуває у контексті автономних систем озброєнь. Міжнародний комітет Червоного Хреста визначає автономну систему озброєнь як «будь-яку систему озброєнь з автономністю у критичних функціях, тобто систему озброєнь, яка може обирати та атакувати цілі без втручання людини» [59, с. 1]. У резолюції ГА ООН A/RES/79/62 «Летальні автономні системи озброєнь» висловлюється стурбованість «можливими негативними наслідками та впливом автономних систем озброєнь на глобальну безпеку та регіональну і міжнародну стабільність, включаючи ризик виникнення гонки озброєнь, зниження порогу конфлікту та розповсюдження, в тому числі до недержавних акторів» [87]. Генеральний секретар ООН Антоніу Гутерреш заявляв, що «автономні машини з владою та дискреційними повноваженнями вибирати цілі та забирати життя без участі людини є політично неприйнятними, морально огидними і повинні бути заборонені міжнародним правом» [61] та закликав до укладання нового міжнародного договору, що встановить конкретні заборони та обмеження на такі системи. Проблематика регулювання автономних систем озброєнь з ШІ, зокрема, питання забезпечення належного людського контролю над критичними функціями таких систем, досліджувалася нами в окремій публікації [62] та залишається одним із найбільш гострих питань сучасного міжнародного права.

Третьою властивістю ШІ є його здатність до автономного функціонування та прийняття рішень без безпосереднього людського втручання. Ця властивість породжує фундаментальні питання щодо юридичної відповідальності за шкоду, заподіяну діями систем ШІ, та контролю над ними. Системи ШІ можуть приймати

рішення на основі складних алгоритмів та великих масивів даних способами, які не завжди можуть бути передбачені або повністю зрозумілі навіть їх розробникам. Непередбачуваність поведінки ШІ створює унікальні виклики для міжнародно-правового регулювання, адже традиційні правові механізми базуються на презумпції передбачуваності та контрольованості дій суб'єктів права. Ця проблема набуває особливої гостроти у контексті так званого ефекту «чорної скриньки». Як зазначає Європейська комісія у Білій книзі про штучний інтелект, «специфічні характеристики багатьох технологій ШІ, включаючи непрозорість (ефект «чорної скриньки»), складність, непредбачуваність та частково автономну поведінку, можуть ускладнити перевірку дотримання та можуть перешкоджати ефективному забезпеченню виконання норм існуючого права ЄС, призначених для захисту основоположних прав» [63].

Проблема «чорної скриньки» полягає у тому, що внутрішня логіка ухвалення рішень системами ШІ є непрозорою навіть для розробників таких систем. В науковій літературі ця проблема розглядається як комплекс взаємопов'язаних питань. Зокрема, як зазначає польський правник Б. Брожек, проблема чорної скриньки фактично є поверхневою, оскільки впливає з накладання чотирьох різних, хоча й взаємопов'язаних питань: проблеми непрозорості (англ. – *opacity problem*), проблеми незвичності (англ. – *strangeness problem*), проблеми непредбачуваності (англ. – *unpredictability problem*) та проблеми обґрунтування (англ. – *justification problem*) [64, с. 427]. Непередбачуваність є особливо проблематичною в контексті правового регулювання, оскільки, як справедливо зазначає Я. Батаї, «непередбачуваність ШІ усуває позитивні ефекти строгої відповідальності. Якщо розробник або користувач ШІ не може передбачити наслідки роботи ШІ *ex ante*, він не може вжити запобіжних заходів щодо заподіяної шкоди» [65, с. 896].

Ця властивість обумовлює потребу у забезпеченні належного рівня контролю та підзвітності систем ШІ, зокрема, через вимоги щодо прозорості та пояснюваності їхніх рішень. Як зазначає Г. Павлідіс, «відсутність пояснюваності штучного інтелекту є однією з перших перешкод, які промисловість та регулятори повинні

подолати, щоб пом'якшити ризики, пов'язані з технологією. Потреба в «зрозумілому ШІ» (англ. – eXplainable AI, XAI) є очевидною в сферах, де підзвітність, етика та справедливість є критичними, таких як охорона здоров'я, кредитний скорінг, поліція та система кримінального правосуддя» [66, с. 293]. Саме ця потреба знаходить нормативне відображення у принципах прозорості та пояснюваності, які належать до ключових засад міжнародно-правового регулювання ШІ.

Четвертою властивістю є динамічний розвиток ШІ, тобто здатність систем до самонавчання та постійної еволюції. На відміну від традиційних технологій, що функціонують відповідно до заздалегідь визначених параметрів, системи ШІ можуть змінювати свою поведінку на основі нового досвіду та даних. Регламент (ЄС) 2024/1689 (Акт про ШІ) визнає цю характеристику, описуючи системи ШІ як такі, що «можуть виявляти адаптивність після розгортання», що означає «здатність системи змінюватися під час використання» [25]. Ця ознака створює особливі труднощі для правового регулювання: система, що відповідає певним технічним стандартам на момент її сертифікації, може суттєво змінити свої характеристики в процесі функціонування, а це обумовлює потребу в таких підходах до регулювання, які б забезпечували не лише попередню оцінку системи, а й постійний її моніторинг протягом усього життєвого циклу. Здатність до самонавчання також загострює питання про збереження належного контролю людини над діяльністю систем ШІ. Як підкреслює МКЧХ, «поняття контролю людиною стало головним питанням у дебатах щодо автономних систем озброєння. Існує широка згода щодо того, що контроль людиною над системами озброєння та застосування сили має бути збережений» [59, с. 5], і ця позиція дедалі більше визнається актуальною і для цивільних застосувань ШІ, про що свідчить його послідовне закріплення у Рекомендації ЮНЕСКО [13, п. 35-36] та Акті про ШІ [25, ст. 14].

П'ята властивість ШІ полягає в його трансформативному впливі на міжнародні відносини та їх правове регулювання. Технології ШІ не просто створюють нові виклики для правового регулювання міжнародних відносин – вони потенційно здатні змінити їх характер і баланс сил між державами, а отже, опосередковано

вплинути й на їх правове регулювання. Як справедливо підкреслюють українські правознавці Д.М. Белов та М.В. Белова, «інтеграція технологій штучного інтелекту у різноманітні аспекти суспільного життя відкриває нові горизонти та водночас створює серйозні виклики для правової системи» [67, с. 818]. Цей вплив уже проявляється: у зміні характеру міжнародної конкуренції (технологічне суперництво провідних держав за лідерство у сфері ШІ); у трансформації механізмів реалізації державного суверенітету в умовах цифровізації (контроль над транскордонними потоками даних); у появі нових форм асиметричних загроз (використання недержавними акторами засобів дезінформації та дипфейків – синтетичних аудіо-, відео- чи зображувальних матеріалів, згенерованих за допомогою ШІ, які реалістично імітують реальних осіб і події з метою введення в оману (англ. – deepfake. – Прим. В. І.)), за яких недержавні актори (зокрема транснаціональні технологічні корпорації, терористичні угруповання) можуть отримати доступ до потужних інструментів впливу на міжнародну безпеку та внутрішні справи держав. Н.О. Попова зазначає, що «недосконалість або навіть відсутність законодавчої бази підвищує рівень та обумовлює появу нових ризиків правопорушень у цій сфері суспільних відносин (тобто у сфері розроблення та застосування ШІ – прим. В. І.), що створює загрозу для національного та міжнародного правопорядку» [3, с. 245].

Здатність ШІ трансформувати міжнародні відносини визнається і на рівні міжнародних організацій. ОЕСР зазначає, що «системи ШІ мають глибокі, далекосяжні та глобальні наслідки, які трансформують суспільства, економічні сектори та світ праці, і, ймовірно, продовжуватимуть це робити в майбутньому» [9]. У Пакті про майбутнє держави-члени ООН визнали, що «цифрові та новітні технології, включаючи штучний інтелект, відіграють значну роль як фактори забезпечення сталого розвитку та кардинально змінюють наш світ», у зв'язку з чим погодились «посилити міжнародне регулювання щодо штучного інтелекту на благо людства» [21]. Українські науковиці Н.В. Камінська та А. Камінська підкреслюють, що «ШІ зумовлює трансформацію міжнародного права, наднаціонального і

національного права. Проте його належне використання потребує ефективного правового регулювання, яке має бути доступним, прозорим, гарантуватиме права і свободи людини, підзвітність і ефективний нагляд» [68, с. 234]. Ця характеристика ставить перед міжнародним правом завдання адаптувати його основоположні принципи – суверенної рівності держав, невтручання у внутрішні справи держави, мирного врегулювання спорів – до якісно нових технологічних реалій.

Сукупність розглянутих властивостей засвідчує унікальність ШІ як об'єкта міжнародно-правового регулювання. Транскордонність, подвійне використання, автономність, здатність до постійної (безперервної) еволюції та трансформативний вплив на міжнародні відносини та їх правове регулювання у своєму поєднанні створюють феномен, який не вписується в усталені категорії міжнародного права (зокрема у традиційний поділ на суб'єктів та об'єктів міжнародно-правового регулювання), розроблені для регулювання відносин між державами та іншими визнаними суб'єктами. Це обумовлює необхідність формування інтегративної дефініції ШІ, яка була б водночас достатньо гнучкою для адаптації до технологічного розвитку та достатньо визначеною для цілей міжнародно-правового регулювання та подальшого правозастосування (дефініція слугує насамперед чіткій ідентифікації об'єкта регулювання). Авторська дефініція, що пропонується нижче, ґрунтується на виявлених властивостях ШІ та результатах порівняльного аналізу підходів до визначення ШІ, закріплених в актах міжнародного права. Питання ж формування спеціального правового режиму для ШІ з урахуванням його унікальних характеристик розглядатиметься у наступному підрозділі.

На основі проведеного аналізу концептуальних підходів та виявлених властивостей штучного інтелекту пропонуємо таку дефініцію ШІ:

Штучний інтелект – це комплексна соціотехнічна система, створена з використанням методів машинного навчання та/або обробки даних чи символічного міркування, що характеризуються:

а) функціональною здатністю до сприйняття та аналізу структурованої й неструктурованої інформації з фізичного або цифрового середовища, її інтерпретації

через процеси інференції (виведення висновків з даних), що виходять за межі простої обробки даних;

б) операційною автономністю різного ступеня при виконанні когнітивних завдань та прийнятті рішень для досягнення явних або неявних цілей без безперервного людського втручання;

в) адаптивністю через здатність до навчання на основі досвіду та модифікації власної поведінки після введення в експлуатацію (розгортання) відповідно до змін середовища функціонування;

г) генеративною спроможністю створювати результати у формі прогнозів, рекомендацій, контенту або рішень, що впливають на фізичне та/або віртуальне середовище, а також на реалізацію прав людини;

д) транскордонним характером застосування та потенціалом подвійного використання;

е) створенням юридично значущих наслідків для прав людини, міжнародного правопорядку та безпеки, що обумовлює необхідність диференційованого міжнародно-правового регулювання залежно від рівня ризиків, які система створює на різних етапах свого життєвого циклу.

Запропонована дефініція втілює інтегративний підхід через поєднання технологічних елементів (методи машинного навчання, обробка даних, символічне міркування), функціональних характеристик (когнітивні завдання, інференція), системного розуміння (соціотехнічна система), антропоцентричного виміру (вплив на права людини), еволюційної природи (адаптивність, життєвий цикл) та правових критеріїв (рівень ризику, характер використання). Дефініція забезпечує технологічну нейтральність визначення завдяки відсутності прив'язки до конкретних технологічних рішень та водночас містить критерії, що можуть застосовуватися для кваліфікації систем як ШІ. Вона інтегрує підходи, закріплені в актах провідних міжнародних організацій, враховує специфіку міжнародно-правового регулювання – транскордонний характер відносин і потребу узгодження інтересів багатьох держав – та відображає необхідність балансу між стимулюванням

інновацій і захистом прав людини. Закладений у дефініції ризикоорієнтований підхід прямо узгоджується з чотирирівневою класифікацією ризиків Акту про ШІ [25].

1.3. Загальний аналіз специфіки міжнародно-правового регулювання штучного інтелекту

Проведене дослідження генези та еволюції міжнародно-правового регулювання штучного інтелекту, а також дослідження поняття і основних властивостей ШІ як об'єкта міжнародно-правового регулювання створили теоретичну основу для розгляду специфіки міжнародно-правового регулювання цього феномену. Як засвідчив проведений у попередніх підрозділах аналіз характеристик ШІ, останній складно повністю підпорядкувати традиційним юридичним категоріям: ШІ поєднує матеріальний (апаратне забезпечення, носії даних) і нематеріальний (алгоритми, моделі) виміри, а також об'єктні риси (ШІ як товар, послуга чи об'єкт права інтелектуальної власності) та функціонально-суб'єктні (здатність до автономного прийняття рішень, що зовні нагадує вольову поведінку суб'єкта), поряд із динамічністю та здатністю до еволюції після введення до експлуатації.

Зазначені особливості ШІ зумовили доктринальну дискусію щодо того, чи слід визнавати за штучним інтелектом самостійний правовий статус суб'єкта права, або ж кваліфікувати його виключно як об'єкт правового регулювання. На думку української вченої Н. М. Оніщенко, «невизначеність ролі (як суб'єкта, так і об'єкта правовідносин) впливає на те, що ШІ породжує, на жаль, не тільки доктринальну суперечливість про перевагу корисності або загроз використання ШІ, а й практичне свідоме або несвідоме блокування, гальмування «легалізації» ШІ-простору в праві» [280, с. 23]. Наведена позиція є показовою для розуміння того, чому формування міжнародно-правового регулювання ШІ відбувається нелінійно та супроводжується тривалим домінуванням «м'якого права» над «твердим»: невизначеність правового

положення ШІ та триваюча у доктрині суперечка щодо визнання його окремим суб'єктом права є не лише суто теоретичною проблемою, а й чинником, що об'єктивно стримує договірну фіксацію зобов'язань держав. Водночас саме ця невизначеність обґрунтовує доцільність обраного у цьому дослідженні інтегративного підходу до визначення ШІ, який, поєднуючи технологічні, функціональні, системні, антропоцентричні та правові критерії, дозволяє кваліфікувати ШІ саме як об'єкт міжнародно-правового регулювання, не зміщуючи його у площину дискусійних концепцій «електронної особи» чи квазісуб'єкта.

Зазначена об'єктна кваліфікація ШІ та поєднання у ньому різнопланових – технічних, економічних і правових – ознак обумовлюють необхідність формування спеціального правового режиму і водночас вимагають інтегративного підходу до правового регулювання ШІ, що поєднує норми різних галузей міжнародного права. Як зазначає українська вчена Л.Г. Фалалєєва щодо взаємодії міжнародного права прав людини та міжнародного гуманітарного права, ці галузі, «маючи обмежені точки дотику і відмінні механізми реалізації, взаємодіють, застосовуються паралельно, взаємодоповнюють одне одного задля належного захисту гарантованих прав» [69, с. 583]. Ця логіка взаємодоповнюваності повною мірою застосовна і до регулювання ШІ: норми міжнародного права прав людини, права міжнародної безпеки, міжнародного економічного права та міжнародного інформаційного права мають функціонувати в режимі комплементарності для забезпечення всебічного та ефективного міжнародно-правового регулювання ШІ. Багатовимірний характер впливу на сучасний міжнародний правопорядок підкреслює й М. О. Медведева, яка зазначає, що використання ШІ кидає виклики «традиційним підходам до міжнародно-правової відповідальності держав та індивідуальної кримінальної й цивільної відповідальності приватних осіб; національній та міжнародній системі захисту прав людини; правилам ведення війни та основним принципам міжнародного гуманітарного права; нормам міжнародного економічного права, зокрема права СОТ; основам здійснення національного та міжнародного правосуддя» [289, с. 185]. Це спостереження підтверджує необхідність формування

спеціального правового режиму, який інтегрує елементи різних галузей міжнародного права, а не вкладається в межі жодної з них.

Специфіка міжнародно-правового регулювання ІІІ визначається передусім тим, що ІІІ, як об'єкт такого регулювання, відрізняється від традиційних об'єктів міжнародного права – територій, природних ресурсів, космічного простору, морських просторів – за трьома ключовими параметрами. По-перше, на відміну від них, ІІІ демонструє різні ступені агентності – від виконання запрограмованих функцій до самостійного прийняття рішень на основі аналізу даних. Це створює нові виклики для міжнародно-правового регулювання відповідальності за наслідки, спричинені діями систем ІІІ, зокрема для визначення суб'єкта, чия поведінка зумовила юридично значущі наслідки. По-друге, визначення відповідальної особи у сфері ІІІ ускладнюється багаторівневою структурою створення та використання системи: розробники алгоритмів, постачальники даних, оператори систем та кінцеві користувачі можуть мати різний ступінь контролю та впливу на результат, тоді як для традиційних об'єктів питання належності та контролю вирішується відносно просто, через концепції суверенітету, власності або юрисдикції. По-третє, міжнародно-правове регулювання ІІІ стикається з безпрецедентними викликами транскордонності та екстериторіальності: якщо традиційні об'єкти мають чітку територіальну прив'язку, то ІІІ функціонує в кіберпросторі, який за своєю природою є транснаціональним, що ставить під сумнів ефективність традиційних принципів територіального верховенства та невтручання у внутрішні справи.

Окремого аналізу заслуговує згадана відмінність – складність визначення особи, яка має нести відповідальність за негативні наслідки роботи системи ІІІ, – оскільки у сфері застосування ІІІ вона набуває якісно нового характеру. Дослідник у галузі філософії техніки А. Маттіас характеризував цю проблему як «розрив відповідальності» (англ. – responsibility gap), визначаючи її як стан, за якого «ніхто не має достатнього контролю над діями машини, щоб мати змогу взяти на себе відповідальність за них» [77, с. 176]. Виникнення такого розриву він пояснював тим, що «правила, за якими вони діють, не є фіксованими під час виробничого процесу, а

можуть змінюватися під час експлуатації машини самою машиною» [77, с. 182]. Розробники алгоритмів можуть стверджувати, що не контролюють результати навчання системи на конкретних даних; постачальники даних – що не впливають на алгоритм; оператори систем – що не мають доступу до «чорної скриньки» алгоритму; кінцеві користувачі – що поклалися на рекомендації системи. Ця багаторівнева структура створює ситуацію «розпорошення» контролю, за якої традиційні механізми притягнення до відповідальності, що передбачають наявність ідентифікованого правопорушника, який здійснює фактичний контроль над поведінкою системи ШІ, виявляються малоприматними. На рівні міжнародного права ця проблема набуває додаткового виміру, коли системи ШІ використовуються державними органами: у такому разі постає питання про можливість розгляду рішень системи ШІ як поведінки держави відповідно до критеріїв, закріплених у Проектах статей Комісії міжнародного права про відповідальність держав за міжнародно-протиправні діяння [70], зокрема щодо поведінки осіб або утворень, уповноважених здійснювати елементи державної влади (ст. 5).

Додаткову складність для міжнародно-правового регулювання створює автономність систем ШІ: чим вищий ступінь автономності системи, тим складніше встановити причинно-наслідковий зв'язок між рішенням людини та результатом діяльності системи. Ця проблема набуває особливої гостроти у контексті автономних систем озброєнь – систем, здатних самостійно здійснювати пошук, вибір та ураження цілі без безпосереднього людського втручання. Такі системи вже не є суто гіпотетичними: до них відносять, зокрема, баражуючі боєприпаси та окремі зразки автономного озброєння, застосування яких задокументовано у сучасних збройних конфліктах, що детально проаналізовано в нашій окремій статті [62]. Не менш значущою ця проблема є у контексті алгоритмічного прийняття рішень у сферах міграційного контролю, соціального забезпечення та кримінального правосуддя, в яких помилки системи ШІ можуть мати серйозні наслідки для прав людини.

Доктринальні підходи до вирішення цієї проблеми варіюються. Американський дослідник права М. Шерер зазначає, що «традиційні методи регулювання – такі як ліцензування продукції, нагляд за дослідженнями і розробками та деліктна відповідальність – видаються особливо непридатними для управління ризиками, пов'язаними з інтелектуальними та автономними машинами» [31, с. 356]. Одним із запропонованих підходів є модель строгої (безвинної) відповідальності: ще у 2017 році Європейський парламент у Резолюції з рекомендаціями щодо норм цивільного права про робототехніку розглядав можливість запровадження «строгої відповідальності (англ. – strict liability. – Прим. В. І.), яка вимагатиме лише доказу шкоди та встановлення причинно-наслідкового зв'язку» [6, п. 53]. Альтернативним підходом є модель розподіленої відповідальності, за якої обов'язки диференціюються між усіма учасниками «життєвого циклу» системи ШІ пропорційно до їх ролі та ступеня контролю. Як буде продемонстровано у Розділі 3, саме цю модель нормативно втілено у Регламенті (ЄС) 2024/1689 (Акті про ШІ) [25], який диференціює обов'язки постачальників, імпортерів, дистриб'юторів та операторів систем ШІ.

Ще одним специфічним викликом для міжнародно-правового регулювання ШІ є проблема, яку ми пропонуємо назвати «темпоральним розривом», – невідповідність між швидкістю технологічного розвитку систем ШІ та повільністю нормотворчих процесів. Сутність цього розриву полягає в розбіжності між темпоральною логікою технологічного розвитку, що вимірюється місяцями, та темпоральною логікою міжнародного нормотворення, що вимірюється роками і десятиліттями. Класична парадигма міжнародного нормотворення, орієнтована на створення стабільних, довготривалих міжнародних договорів, не пристосована до регулювання технології, що зазнає фундаментальних трансформацій у проміжках часу, коротших за тривалість дипломатичних переговорів. Ця проблема має структурний характер: вона не зумовлена недоліками конкретної організації чи процедури, а впливає із самої природи ШІ як динамічного об'єкта, здатного до еволюції після розгортання. Показово, що споріднену асинхронність вітчизняна

доктрина фіксує її на національному рівні: як зазначає М. І. Суржинський, «стрімке вдосконалення систем ШІ неминуче руйнує традиційні ритми правотворчості: законодавчі цикли (концептуалізація, консультації, імплементація, перегляд) часто виявляються повільнішими за темп технологічних інновацій. У результаті формується «регуляторне відставання» – часовий розрив між фактичним ризик-профілем технологій і наявними правовими приписами» [284, с. 333]. Запропонована нами категорія є відмінною за акцентом: якщо «регуляторне відставання» описує результат – невідповідність чинних приписів актуальному ризик-профілю технологій, – то «темпоральний розрив» фіксує структурну причину цього результату – розбіжність самих темпоральних логік технологічного розвитку та міжнародного нормотворення, яка не усувається разовим оновленням норм і потребує зміни моделі нормотворення. Структурність розриву посилюється її епістемічною невизначеністю: як зауважує М. Шерер, «експерти з ШІ сьогодні, таким чином, можуть мати труднощі з оцінкою ризиків, пов'язаних з ШІ завтра» [31, с. 385]. Темпоральний розрив має при цьому не лише описовий, а й нормативний вимір. Як слушно зауважує О. В. Костенко, «критичним є питання темпоральності: чи можна вважати «вичікування» нейтральною позицією. Необхідно вважати, що нейтральності тут немає, оскільки регуляторна пауза фактично дозволяє масштабувати практики, здатні спричиняти дискримінацію, необґрунтоване втручання у приватність, обмеження доступу до соціальних послуг або непрозорі санкції, а отже – акумулювати шкоду до того, як будуть створені процедурні бар'єри» [283, с. 172]. Відтак повільність міжнародного нормотворення не є ціннісно нейтральною: вона сама стає чинником ризику, що додатково обґрунтовує як запобіжний підхід, так і потребу в нових моделях правового регулювання, здатних скоротити тривалість нормотворчих циклів.

Зазначимо, що міжнародно-правове регулювання інших технологічних сфер стикалося зі схожими викликами, що дозволило виробити кілька підходів до адаптації правового регулювання до технологічних змін. Конструкція рамкових конвенцій з протоколами (наприклад, Віденська конвенція про захист озонового

шару 1985 року з Монреальським протоколом 1987 року [75]), дозволяє встановити загальні принципи у рамковій конвенції, а конкретні зобов'язання деталізувати у протоколах, що можуть переглядатися оперативніше. Прийом нормотворчої техніки, відомий як «sunset clauses» (норми із визначеним терміном дії або вбудованим обов'язком перегляду), створює необхідність періодичної переоцінки доцільності регуляторних приписів – елементи цього підходу вже застосовано у Регламенті (ЄС) 2024/1689 (Акті про ШІ), ст. 112 якого встановлює обов'язок Комісії здійснювати періодичну оцінку та перегляд Регламенту [25]. «Регуляторні пісочниці» (англ. – regulatory sandboxes), що набули поширення в регулюванні фінансових технологій та знайшли нормативне закріплення у ст. 57–60 Акту про ШІ [25], дозволяють тестувати інноваційні рішення в контрольованому середовищі до прийняття остаточних регуляторних рішень. Так, Акт про ШІ зобов'язує держави-члени ЄС створити такі пісочниці на національному рівні для випробування систем ШІ під наглядом компетентних органів [25, ст. 57].

Водночас ШІ створює додаткову складність, відсутню в інших сферах суспільних відносин. На відміну від більшості інших технологій, об'єкт правового регулювання яких залишається відносно стабільним після створення, системи ШІ продовжують еволюціонувати після розгортання внаслідок процесів машинного навчання. Система, що відповідала регуляторним вимогам на момент введення в експлуатацію, може суттєво змінити свою поведінку після навчання на нових даних, що створює для міжнародно-правового регулювання проблему «рухомої цілі» (англ. – moving target). Саме здатність системи змінювати свою поведінку вже після розгортання (проблема «рухомої цілі») й обумовлює потребу у таких підходах до правового регулювання, які забезпечували б не лише адаптацію норм до появи нових типів систем, а й постійний моніторинг уже розгорнутих систем ШІ протягом усього їх життєвого циклу. Теоретичне осмислення цієї проблеми та обґрунтування відповідної моделі правового регулювання буде здійснено у Розділі 2 на основі емпіричного аналізу актів універсального та міжрегіонального рівнів.

Міжнародно-правове регулювання ШІ також характеризується фрагментацією регуляторних підходів, що відображає як об'єктивну складність предмета регулювання, так і загальну тенденцію до фрагментації міжнародного права. На структурний характер цієї тенденції вказують, зокрема, П. Сайхон, М. Маас і Л. Кемп, які досліджують архітектуру міжнародно-правового регулювання ШІ крізь призму його фрагментації [148]. На універсальному рівні, незважаючи на активну діяльність міжнародних організацій, відсутній єдиний міжнародно-правовий акт, навіть рамкового характеру, який встановлював би обов'язкові для всіх держав принципи та норми регулювання ШІ. Замість цього спостерігається мозаїка різномірних ініціатив – від резолюцій Генеральної Асамблеї ООН та рекомендацій ЮНЕСКО до стратегічних документів спеціалізованих установ ООН, – кожна з яких фокусується на окремих аспектах проблеми, не створюючи цілісної регуляторної системи з юридично обов'язковою силою. Як справедливо зазначають науковці Е. Кляйн та С. Патрік, «нинішня однодушність щодо узгодження ШІ з людськими цінностями приховує фрагментацію, яка вже характеризує національні регуляторні підходи» [46, с. 12]. На регіональному рівні фрагментація проявляється в диверсифікації регуляторних моделей, що відображають різні правові традиції, політичні системи та економічні пріоритети: від детального нормативного регулювання ЄС, яке охоплює різні сектори застосування ШІ, до добровільних керівних принципів (настанов) АСЕАН та декларативних документів інших регіональних організацій, які будуть детально проаналізовані у Розділі 3 цього дослідження. Проблема фрагментації посилюється відсутністю координації між різними рівнями регулювання, що створює ризик «регуляторного арбітражу», при якому розробники систем ШІ можуть обирати юрисдикції з найменш обтяжливими вимогами, підриваючи ефективність окремих регуляторних режимів.

Водночас, попри фрагментацію регуляторних підходів, спостерігається помітна конвергенція принципів міжнародно-правового регулювання ШІ – тобто поступове формування спільного кола принципів, на яких ґрунтується таке регулювання. Дослідження Дж. Фелд, Н. Ахтена, Х. Хіллігос, А. К. Надя та М.

Шрікумара виявило високий ступінь узгодженості міжнародно-правових актів навколо восьми ключових груп принципів: прозорість та пояснюваність, підзвітність та відповідальність, справедливість та недискримінація, приватність та захист даних, безпека та захищеність, людський контроль, професійна відповідальність та просування людських цінностей [80, с. 21]. Особливо показовим є те, що 64% проаналізованих науковцями джерел містять посилання на права людини [80, с. 6], що свідчить про формування консенсусу щодо того, що регулювання ШІ повинно ґрунтуватися на міжнародних стандартах прав людини, закріплених у Загальній декларації прав людини (1948), Міжнародному пакті про громадянські та політичні права (1966) та Міжнародному пакті про економічні, соціальні та культурні права (1966). Як підтверджує дослідник Н. Корреа, ця конвергенція не є випадковою, а відображає перетворення «принципізму» (англ. – *principilism*) – методологічного підходу, що ґрунтується на формулюванні високорівневих етичних принципів як основи регулювання, – на «резонуючу основу для більшості різноманітних етичних рекомендацій, політичних рамок, правових орієнтирів, кодексів поведінки» [134, с. 4].

Разом з тим конвергенція принципів міжнародно-правового регулювання ШІ не усуває відмінностей у практиці їх реалізації. Як засвідчує спільне дослідження Г. Робертса, Е. Гайн, М. Таддео та Л. Флоріді, присвячене політиці управління ШІ, «високорівневі принципи ... є розпливчастими та допускають різні ідеологічні позиції» [143, с. 1277]. Автори наводять показовий приклад: принцип справедливості ШІ, підтриманий усіма членами G20 стосовно технологій розпізнавання обличчя, тлумачиться в ЄС як заборона цих технологій, тоді як у Китаї вони вважаються допустимими в інтересах соціальної стабільності [143, с. 1277]. Зокрема, зберігається дивергенція у підходах до регулювання воєнного використання ШІ, соціального скорингу та масового спостереження – сфер, в яких національні інтереси та ціннісні орієнтації держав розходяться найбільш суттєво. Цей аналіз дозволяє констатувати діалектичну єдність двох тенденцій як визначальну характеристику міжнародно-правового регулювання ШІ:

запровадження спільних принципів і цінностей у регулюванні ШІ, з одного боку, та фрагментацію практики їх реалізації – на рівні конкретних регуляторних моделей і механізмів імплементації, – з іншого. Ця закономірність буде емпірично підтверджена у Розділі 2 на прикладі аналізу резолюцій ГА ООН та актів спеціалізованих установ, а у Розділі 3 – на прикладі регіональних ініціатив.

Виявлена закономірність – поєднання фрагментації регуляторних підходів зі стійкою конвергенцією базових принципів – має й суттєве доктринальне підґрунтя у теорії права. Як зазначає українська вчена Н. М. Оніщенко, «не заперечуючи того, що правове регулювання простору штучного інтелекту може й повинно мати певні особливості, акцентуємо увагу на неприйнятності його відокремлення від загальноприйнятих канонів, норм і механізмів правового регулювання» [280, с. 28]. Ця теза становить важливий концептуальний орієнтир для оцінки специфіки міжнародно-правового регулювання ШІ: попри визнання його *sui generis* природи, регуляторний режим ШІ не може формуватися ізольовано від основоположних засад, цінностей та інституційних механізмів міжнародного права. Саме тому спостережувана конвергенція навколо принципів людиноцентричного підходу, поваги до прав людини, прозорості, підзвітності та недискримінації не є випадковою – вона відображає необхідність відповідності нового регуляторного режиму основоположним принципам міжнародного права, а також нормам міжнародного права прав людини і права міжнародних договорів. Відхід від цих засад, на що цілком обґрунтовано вказує Н. М. Оніщенко, може спричиняти фрагментарне або ізольоване сприйняття правового режиму ШІ, що загрожує цілісності міжнародно-правового регулювання в умовах цифрової трансформації [280].

Поєднання фрагментації – на рівні регуляторних підходів, інституційних механізмів та практики правової реалізації – з конвергенцією принципів і цінностей визначає особливу роль «м'якого права» у міжнародно-правовому регулюванні ШІ. Домінування «м'якого права» на сучасному етапі міжнародно-правового регулювання ШІ обумовлено кількома об'єктивними чинниками. По-перше, динамічність технологічного розвитку, що проявляється у виявленому

«темпоральному розриві», робить традиційні процеси створення юридично обов'язкових норм занадто повільними для адекватного реагування на нові виклики. По-друге, складність досягнення консенсусу між державами щодо міжнародно-правових зобов'язань у сфері, яка має стратегічне значення для національної конкурентоспроможності та безпеки, зумовлює обережність у прийнятті обов'язкових міжнародно-правових актів; «м'яке право» дозволяє частково подолати цю складність, забезпечуючи попереднє узгодження позицій держав без прийняття ними юридично обов'язкових зобов'язань. По-третє, невизначеність щодо довгострокових наслідків розвитку ШІ – аналогічна до тієї, що спостерігалася на ранніх етапах регулювання біотехнологій, – створює потребу в таких формах правового регулювання, що допускають оперативний перегляд та адаптацію до технологічних змін.

Спостережувана динаміка міжнародно-правового регулювання ШІ полягає у поступовому русі від переважно «м'якого права» до появи перших юридично обов'язкових («твердих») норм, найвиразніше виявленому на регіональному рівні (Розділ 3). Разом з тим цей перехід не означає повного заміщення одного типу норм іншим. Навпаки, обидва типи продовжують співіснувати та взаємодіяти, утворюючи, за влучним визначенням Г. Робертса, М. Таддео та Л. Флоріді, не бінарну дихотомію «обов'язкове – необов'язкове», а «континуум за трьома вимірами – ступенем зобов'язання, точності та делегування» [150, с. 8]. «М'яке право» зберігає свою роль у сферах, де консенсус щодо юридично обов'язкового регулювання ще не досягнутий, або де потрібна гнучкість для адаптації до технологічних змін. Як буде детально обґрунтовано у Розділі 2 на підставі емпіричного аналізу актів універсального та міжрегіонального рівнів, «м'яке право» у сфері регулювання ШІ виконує низку специфічних функцій, що мають критичне значення для поступового формування цілісного міжнародно-правового режиму. Водночас саме лише «м'яке право» є недостатнім для ефективного регулювання технологій із такими потужними та потенційно необоротними наслідками, як ШІ, що обумовлює

необхідність його доповнення юридично обов'язковими актами; як засвідчить аналіз у Розділі 3, перші такі акти формуються наразі саме на регіональному рівні.

Специфіка міжнародно-правового регулювання ШІ проявляється також у формуванні ризикоорієнтованого підходу як одного з ключових регуляторних підходів. Цей підхід передбачає диференціацію регуляторних вимог залежно від рівня ризику, який створює конкретна система ШІ. З теоретичної точки зору, ризикоорієнтований підхід відображає застосування принципу пропорційності в міжнародному праві: обсяг регуляторних вимог повинен бути пропорційним рівню потенційних ризиків для прав людини та суспільних інтересів. Водночас він поєднує превентивний принцип, визнаний у міжнародному екологічному праві та закріплений, зокрема, у Принципі 15 Декларації Ріо-де-Жанейро з навколишнього середовища та розвитку 1992 року [76], з необхідністю збереження простору для технологічного розвитку та інновацій. Саме ця здатність до балансування пояснює поширення ризикоорієнтованого підходу: він дозволяє уникнути як надмірного регулювання, що може загальмувати інноваційний розвиток, так і неповного регулювання, що залишає незахищеними фундаментальні права людини. Конкретні прояви ризикоорієнтованого підходу в регіональних актах, зокрема чотирирівнева класифікація ризиків у Регламенті (ЄС) 2024/1689 (Акті про ШІ) та градуїований підхід Рамкової конвенції Ради Європи про ШІ, будуть детально досліджені у Розділі 3.

Специфічною рисою міжнародно-правового регулювання ШІ є також участь багатьох зацікавлених сторін у процесі нормотворення, за якого нормотворчий процес у цій сфері виходить за межі класичної міждержавної моделі та залучає широке коло недержавних акторів. Як підкреслює М. Чінен, «у сфері регулювання ШІ формується особлива система правил, що включає комбінацію «твердого» та «м'якого права», які виникають зі складних взаємодій декількох джерел цього права та акторів, які їх пропонують: приватне впорядкування компаніями та асоціаціями; розробники та науковці у сфері ШІ; національні закони та регуляції, прийняті державами; корпус міжнародного права як такий; міжнародні організації;

громадянське суспільство» [81, с. 1]. Ця обставина має принципове значення для міжнародно-правового регулювання ІІІ, оскільки рішення щодо ключових параметрів систем ІІІ – їх архітектури, складу навчальних даних, налаштувань безпеки та умов розгортання – фактично ухвалюються переважно приватними компаніями, а не державами чи міжнародними організаціями і здебільшого поза межами приписів міжнародного права.

Залучення приватних суб'єктів до фактичного ухвалення регуляторно значущих рішень не є абсолютно безпрецедентним для міжнародного права. Урядування в мережі Інтернет (англ. Internet governance) здійснюється за моделлю, що залучає різні категорії зацікавлених сторін, діяльність яких має транснаціональний характер: від Інтернет-корпорації з присвоєння імен і номерів (ICANN) – некомерційної організації приватного права, що виконує функції публічного характеру у сфері координації системи доменних імен та адресного простору Інтернету, – до Форуму з управління Інтернетом (IGF) як майданчика діалогу зацікавлених сторін під егідою ООН; жодна з цих структур не має класичної міждержавної природи. Стандартизація у сфері інформаційних технологій здійснюється переважно через міжнародні неурядові організації зі стандартизації, як-от ISO та IEC, у діяльності яких провідну роль відіграють представники промисловості та бізнесу. Водночас у сфері ІІІ залучення недержавних акторів набуває якісно нового масштабу. Й. Таллберг характеризує архітектуру регулювання ІІІ, що формується, як «режимний комплекс – структуру частково перекриваючих та різноманітних механізмів управління без чітко визначеної центральної інституції або ієрархії» [146, с. 3], причому недержавні актори часто мають більший технічний потенціал та ресурси, ніж більшість держав. Ця асиметрія ресурсів та експертизи породжує проблему легітимності та підзвітності: рішення, що впливають на фундаментальні права мільйонів людей, ухвалюються вузьким колом транснаціональних корпорацій (приватних юридичних осіб), переважно зосереджених у кількох юрисдикціях, які де-факто виконують регуляторні функції без відповідного публічно-правового мандату та суспільного контролю. Водночас,

як зазначають Г. Робертс, М. Таддео та Л. Флоріді, «приватні інституції є ефективними в управлінні питаннями, пов'язаними з технологічною та транзакційною взаємопов'язаністю, оскільки це відповідає їхнім фінансовим інтересам», проте «вони позбавлені стимулу значущим чином вирішувати питання, пов'язані з екологічними збитками чи безпекою ШІ, що свідчить про необхідність втручання публічного сектору» [150, с. 9]. Саме тому міжнародно-правове регулювання ШІ потребує таких правових та організаційних форм, які б забезпечували залучення недержавних акторів при збереженні провідної ролі держав та міжнародних міжурядових організацій як гарантів забезпечення публічного інтересу.

Нарешті, специфічним регуляторним підходом, що набуває дедалі більшого значення в міжнародно-правовому регулюванні ШІ, є підхід на основі «життєвого циклу» системи ШІ. На відміну від більшості традиційних об'єктів міжнародного права, правовий статус та характеристики яких залишаються відносно стабільними після їх виникнення чи створення, системи ШІ проходять послідовні етапи, кожен з яких формує специфічні правові виклики та потребує диференційованих нормативних вимог. «Життєвий цикл» системи ШІ охоплює етапи проєктування, збору та підготовки навчальних даних, навчання моделі, тестування та валідації, розгортання та введення в експлуатацію, моніторингу та підтримки під час функціонування, а також виведення з експлуатації. Концепція «життєвого циклу» знайшла нормативне закріплення у відповідних міжнародно-правових актах. Рекомендація ЮНЕСКО з етики ШІ вказує, що «відповідна оцінка впливу повинна враховувати весь життєвий цикл системи ШІ» [13]. Оновлене визначення ОЕСР 2023 року характеризує системи ШІ через призму їхнього функціонування та підкреслює, що «різні системи ШІ відрізняються за рівнями автономності та адаптивності після розгортання» [9], що імпліцитно визнає зміну характеристик системи протягом її життєвого циклу. Значення концепції «життєвого циклу» для міжнародно-правового регулювання ШІ полягає у тому, що вона дозволяє розширити регуляторне охоплення за межі моменту введення системи в експлуатацію, поширюючи його на

всі етапи існування системи – від проектування до виведення з експлуатації. Це є особливо важливим з огляду на здатність систем ШІ суттєво змінювати свою поведінку після розгортання внаслідок безперервного навчання на нових даних. Підхід на основі «життєвого циклу» створює передумови для диференціації регуляторних вимог залежно від етапу, розподілу обов'язків між різними учасниками процесу створення та використання системи ШІ пропорційно до їхньої ролі та ступеня контролю, а також для забезпечення безперервного моніторингу, документування та підзвітності протягом усього періоду функціонування системи. Як буде продемонстровано у Розділі 2, концепція «життєвого циклу» систематично закріплюється в резолюціях ГА ООН, стаючи одним із базових регуляторних підходів на універсальному рівні міжнародно-правового регулювання ШІ.

Виявлена у цьому підрозділі сукупність специфічних рис міжнародно-правового регулювання ШІ обумовлює необхідність визнання ШІ особливим об'єктом міжнародно-правового регулювання, що потребує формування спеціального правового режиму *sui generis*. Ця концепція дозволяє визнати унікальність феномену ШІ без необхідності підводити його під існуючі категорії об'єктів або суб'єктів міжнародного права і водночас обґрунтовує формування такого правового режиму, який був би здатний враховувати специфіку різних типів ШІ та сфер його застосування, а також еволюціонувати разом із розвитком технології. Формування нових правових категорій при цьому не означає повного відриву від існуючої системи міжнародного права, а становить її еволюційний розвиток, який спирається на фундаментальні принципи – суверенну рівність держав, добросовісне виконання зобов'язань, повагу до прав людини – та адаптує їх до реалій цифрової епохи.

Міжнародне право вже має прецеденти створення спеціальних правових режимів *sui generis* для унікальних об'єктів, аналіз яких свідчить на користь застосування такого підходу до ШІ. Договір про Антарктику 1959 р. [71] створив режим «заморожування» територіальних претензій заради спільного блага, продемонструвавши можливість досягнення міжнародного консенсусу щодо

об'єкта, стосовно якого існують конкуруючі інтереси держав; антарктичний режим також ґрунтується на принципі відкритості та обміну інформацією, що перегукується з принципами прозорості у сфері ШІ. Договір про принципи діяльності держав з дослідження та використання космічного простору 1967 року [72] запровадив концепцію космосу як надбання усього людства, згідно з якою дослідження та використання космічного простору здійснюються «на благо та в інтересах усіх країн, незалежно від ступеня їх економічного чи наукового розвитку» [72], – ця ідея має аналогію з принципом подолання цифрового розриву та інклюзивного доступу до технологій ШІ, який, як буде продемонстровано у Розділі 2, послідовно закріплюється в резолюціях ГА ООН. Конвенція ООН з морського права 1982 року [73] запровадила градуїований режим, що диференціює правовий статус та обсяг прав і обов'язків залежно від зони, – модель, що є структурно подібною до ризикоорієнтованого підходу в регулюванні ШІ. Картахенський протокол про біобезпеку 2000 року [74] закріпив принцип обережності як основу для прийняття регуляторних рішень в умовах невизначеності щодо довгострокових наслідків технологій, що є актуальним і для регулювання систем ШІ, здатних до самонавчання. Монреальський протокол про речовини, що руйнують озоновий шар, 1987 р. [75] продемонстрував ефективність механізму регулярного перегляду та оновлення нормативних приписів у відповідь на нові наукові дані (зокрема ст. 6 та ст. 2 Протоколу, що передбачають періодичну переоцінку регуляторних заходів), – модель, яка може слугувати відповіддю на виявлений «темпоральний розрив». Сукупність виявлених моделей – «заморожування» конкуруючих інтересів, інклюзивний доступ, градуїований режим, принцип обережності та регулярний перегляд норм – свідчить про те, що міжнародне право має релевантний досвід, який може бути використаний при формуванні спеціального правового режиму *sui generis* для ШІ. Такий режим має ґрунтуватися на адаптації та операціоналізації принципів прозорості, підзвітності, людського контролю, справедливості та недискримінації до специфіки функціонування систем ШІ, а також на створенні механізмів розподілу відповідальності між різними учасниками «життєвого циклу» цих систем.

Виявлена у цьому підрозділі специфіка міжнародно-правового регулювання ШІ та обґрунтована необхідність формування правового режиму *sui generis* окреслюють передумови для формування універсальної конвенції зі штучного інтелекту. Реалізація цієї перспективи залежить від здатності держав перевести виявлену конвергенцію принципів у площину юридично обов'язкових зобов'язань, забезпечити баланс між захистом прав людини та інтересами технологічного розвитку і національної безпеки, а також закласти у правовий режим достатню гнучкість та адаптивність до технологічних змін. Рамкова конвенція Ради Європи про штучний інтелект та права людини, демократію та верховенство права, 2024 р. [24], будучи відкритою для приєднання держав з різних регіонів світу, створює потенціал для поступового розширення міжнародного консенсусу за межі європейського регіону, хоча об'єктивні перешкоди – геополітична конкуренція між провідними технологічними державами, різниця у правових традиціях та цінностях, а також економічні інтереси держав та корпорацій – свідчать про тривалість і поступовість цього процесу. Детальний аналіз правової природи та змісту окремих міжнародно-правових актів на універсальному та міжрегіональному рівнях, включаючи емпіричне обґрунтування виявлених у цьому розділі теоретичних закономірностей, буде здійснено у Розділі 2, а дослідження регіонального рівня, на якому, зокрема, відбувається перехід від «м'якого» до «твердого права», – у Розділі 3 дисертаційного дослідження.

Висновки до Розділу 1

Проведене дослідження характеристик ШІ як об'єкта міжнародно-правового регулювання дозволяє сформулювати такі узагальнюючі висновки.

Формування концепції штучного інтелекту охоплює понад сім десятиліть – від Дартмутської конференції 1956 року, для проведення якої Дж. Маккарті запропонував сам термін «штучний інтелект», до етапу інтенсивної нормотворчості, що триває нині. У міжнародно-правовому дискурсі це формування відбувалося

нелінійно: ШІ пройшов шлях від суто технічного феномену через етичні принципи та інструменти «м'якого права» до перших юридично обов'язкових актів у цій сфері, ухвалених 2024 року, які, однак, не є кінцевою точкою цього процесу, а знаменують якісний перехід до нового етапу, для якого характерні юридична обов'язковість приписів, диференціація зобов'язань між учасниками життєвого циклу систем ШІ та запровадження механізмів нагляду й відповідальності. Систематичні спроби такого регулювання розпочалися лише в середині 2010-х років, що зумовлено сукупністю технологічних, соціально-політичних та геополітичних чинників, зокрема проривом у галузі глибокого навчання та проникненням систем ШІ у критичні сфери суспільного життя.

Відсутність універсального визначення ШІ в міжнародно-правових актах та доктрині породжує невизначеність меж предмета регулювання та ускладнює узгодження зобов'язань держав. Подоланню цієї невизначеності слугує здійснений порівняльний аналіз, який дозволив виокремити сім концептуальних підходів до визначення ШІ (технологічний, антропоцентричний, функціональний, системний, еволюційний, цілеорієнтований та інтегративний) і встановити, що, попри їх множинність, спостерігається конвергенція характеристик систем ШІ: здатності до обробки даних і виведення висновків, автономності функціонування, адаптивності та здатності до навчання, цілеспрямованості діяльності, взаємодії з оточуючим середовищем. Наявний консенсус щодо цих характеристик сприяє зближенню підходів як у межах «м'якого права», так і в юридично обов'язкових актах, тоді як відмінності у їх застосуванні зумовлені об'єктивними чинниками – різними регуляторними цілями актів, особливостями правових систем держав та специфікою сфер застосування норм.

Штучному інтелекту як об'єкту міжнародно-правового регулювання притаманні п'ять властивостей, що детермінують особливості відповідного правового режиму: транскордонний характер, що створює виклик для системи міжнародного права, побудованої на принципі територіального суверенітету держав; потенціал подвійного використання – придатність однієї технології як для

цивільних, так і для воєнних цілей; автономність різного ступеня – від повної залежності від людських інструкцій до високого рівня самостійності у прийнятті рішень, що ускладнює притягнення до відповідальності та визначення суб'єкта, чия поведінка спричинила юридично значущі наслідки; здатність до еволюції та самовдосконалення систем машинного навчання, що породжує проблему «рухомої цілі» для правового регулювання; нарешті, суттєвий вплив на міжнародні відносини та їх правове регулювання, який викликає потребу у пристосуванні низки фундаментальних принципів міжнародного права – суверенної рівності держав, невтручання у внутрішні справи та мирного врегулювання спорів – до новітньої практики застосування ШІ.

Специфіка міжнародно-правового регулювання ШІ зумовлена низкою проблем, на подолання яких воно спрямоване, та виражається у відповідних регуляторних підходах. До таких проблем належать: складність визначення відповідальної особи та «розрив відповідальності», що виникає внаслідок «розпорошення» контролю між учасниками створення та використання систем ШІ; «темпоральний розрив» – невідповідність між швидкістю технологічного розвитку та повільністю міжнародного нормотворення, що потребує нових моделей правового регулювання; діалектична єдність фрагментації регуляторних підходів до ШІ та конвергенції фундаментальних принципів його регулювання. Відповіддю на ці виклики стали утвердження підходу за участю багатьох зацікавлених сторін, за якого нормотворчий процес виходить за межі класичної міждержавної моделі, та нормативне закріплення підходу на основі «життєвого циклу» системи ШІ як основи для диференціації регуляторних вимог і розподілу обов'язків між учасниками.

Штучний інтелект не набув у міжнародному праві самостійного правового статусу. Спроби конструювати для систем ШІ окрему категорію міжнародної правосуб'єктності не отримали нормативної підтримки в жодному чинному юридично обов'язковому акті; як Рамкова конвенція Ради Європи про штучний інтелект та права людини, демократію і верховенство права 2024 р., так і Регламент (ЄС) 2024/1689 (Акт про ШІ) послідовно виходять із кваліфікації ШІ як об'єкта

регулювання, зобов'язання щодо якого розподіляються між державами, постачальниками та розгортачами систем ШІ. Відтак міжнародно-правове положення ШІ вичерпно описується категорією об'єкта регулювання, а не суб'єкта.

Сукупність специфічних рис міжнародно-правового регулювання ШІ обумовлює необхідність визнання штучного інтелекту об'єктом міжнародно-правового регулювання *sui generis*, що потребує формування спеціального правового режиму. Унікальність правової природи ШІ – поєднання матеріальних і нематеріальних складових, об'єктних і суб'єктних рис, динамічності та еволюційності – та його відмінності від традиційних об'єктів ускладнюють ефективне застосування наявних правових категорій, механізмів та інструментів регулювання. Водночас міжнародне право вже має прецеденти створення спеціальних режимів *sui generis* для унікальних об'єктів: для Антарктики (Договір про Антарктику 1959 р.), космічного простору (Договір про космос 1967 р.) та дна морів і океанів за межами національної юрисдикції (Конвенція ООН з морського права 1982 р.), – інструменти яких можуть бути застосовані до правового регулювання ШІ.

На основі аналізу різних концептуальних підходів до визначення ШІ та сутнісних характеристик сформульовано авторську дефініцію ШІ як комплексної соціотехнічної системи, що характеризується функціональною здатністю до інференції (виведення висновків з даних), операційною автономністю різного ступеня, адаптивністю через здатність до навчання та модифікації поведінки після розгортання, генеративною спроможністю створювати прогнози, рекомендації, контент або рішення, транскордонним характером застосування та потенціалом подвійного використання, а також створенням юридично значущих наслідків для сфери прав людини, міжнародного правопорядку та безпеки. Запропонована дефініція втілює інтегративний підхід, поєднуючи технологічні, функціональні, системні, антропоцентричні, еволюційні та правові елементи, та забезпечує технологічну нейтральність визначення і його придатність для міжнародно-правових актів різного рівня.

Міжнародно-правове регулювання ШІ еволюціонувало від етичних принципів до юридично обов'язкових норм. «М'яке право» продовжує відігравати ключову роль у цьому регулюванні навіть після появи перших актів «твердого права»: воно забезпечує його гнучкість та адаптивність до швидких технологічних змін, слугує випробувальним майданчиком для нових норм і сприяє формуванню міжнародного консенсусу в умовах геополітичної фрагментації. Обидва типи норм не заміщують один одного, а співіснують і взаємодіють, утворюючи «континуум», що відображає закономірну еволюцію правового регулювання нових технологічних феноменів.

Для міжнародно-правового регулювання ШІ характерна конвергенція його принципів, попри фрагментацію підходів до його здійснення. Міжнародно-правові акти щодо відповідної сфери відносин демонструють узгодженість навколо восьми ключових груп принципів: прозорості та пояснюваності, підзвітності та відповідальності, справедливості та недискримінації, приватності та захисту даних, безпеки та захищеності, людського контролю, професійної відповідальності та просування людських цінностей. Переважна більшість аналізованих актів вказує на необхідність дотримання прав людини, що свідчить про формування консенсусу щодо того, що регулювання ШІ має ґрунтуватися на міжнародних стандартах прав людини.

Специфіка міжнародно-правового регулювання ШІ та необхідність формування правового режиму *sui generis* окреслюють передумови для укладення універсальної конвенції зі штучного інтелекту. Реалізація цієї ідеї залежить від здатності держав перевести конвергенцію принципів у площину юридично обов'язкових зобов'язань, забезпечити баланс між захистом прав людини та інтересами технологічного розвитку і національної безпеки, а також закласти у правовий режим достатню гнучкість та адаптивність до технологічних змін. Водночас об'єктивні перешкоди – геополітична конкуренція між провідними технологічними державами, відмінності у правових традиціях та цінностях, фрагментація регуляторних підходів – свідчать про те, що процес формування такої конвенції може бути тривалим і потребуватиме залучення до нормотворення не

лише держав, а й міжнародних організацій, суб'єктів приватного права, громадянського суспільства та наукових установ.

РОЗДІЛ 2. МІЖНАРОДНО-ПРАВОВЕ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ НА УНІВЕРСАЛЬНОМУ ТА МІЖРЕГІОНАЛЬНОМУ РІВНЯХ

2.1. Напрямки роботи ООН щодо правового регулювання ШІ в інтересах забезпечення сталого розвитку

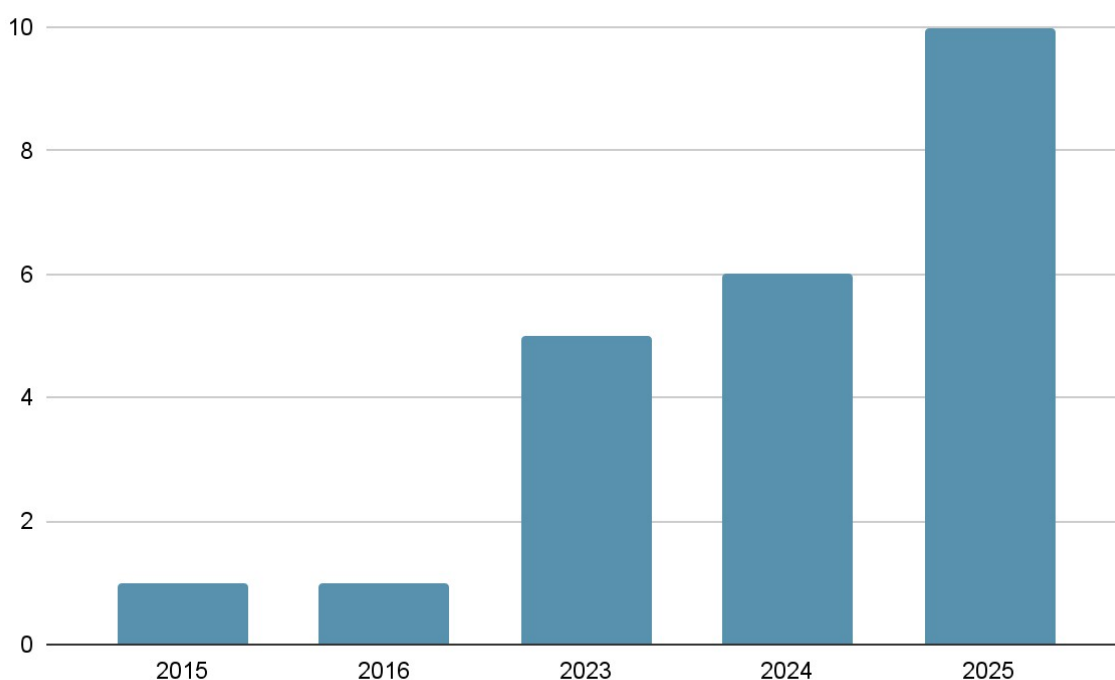
Суспільні відносини в сфері використання ШІ є новим об'єктом для правового регулювання на всіх рівнях – універсальному, міжрегіональному, регіональному та національному. Його складність і багатовимірність зумовлена взаємодією різноманітних джерел права та різноманітних за своїм правовим статусом зацікавлених сторін. Як підкреслює М. Чінен, «у сфері регулювання ШІ формується особлива система правил, що включає комбінацію твердого та м'якого права, які виникають зі складних взаємодій декількох джерел цього права та акторів, які їх пропонують: приватне впорядкування компаніями та асоціаціями; розробники та науковці у сфері ШІ; національні закони та регуляції, прийняті державами; корпус міжнародного права як такий; міжнародні організації; громадянське суспільство» [81, с. 1]. За таких умов, Організація Об'єднаних Націй становить ключову платформу для досягнення консенсусу між державами-членами щодо базових принципів регулювання технологій ШІ в інтересах реалізації Порядку денного сталого розвитку до 2030 року. Причому особливого значення набувають резолюції Генеральної Асамблеї ООН, які, за характеристикою О.В. Кресіна, наукова спільнота визнає «проектами майбутнього права (*de lege ferenda*), м'яким правом, допоміжним джерелом міжнародного права, які діють у сірій зоні між правом і політикою» [22, с. 54]. Вибір резолюцій ГА ООН як основного предмета аналізу у цьому підрозділі зумовлений тим, що ГА ООН є єдиним органом ООН в якому представлені всі 193 держави-члени, що надає їй резолюціям найвищий ступінь політичної легітимності. Крім того, під її егідою формуються загальні принципи та концептуальні рамки регулювання ШІ, тоді як діяльність спеціалізованих установ

ООН, що конкретизує ці принципи через секторальні акти та стандарти, є предметом окремого аналізу в наступному підрозділі.

Хронологічні рамки дослідження обумовлені предметом аналізу: резолюції ГА ООН, прийняті до 2015 року, не містять жодних згадок чи посилань на штучний інтелект як об'єкт регулювання. Початковою точкою відліку обрано 2015 рік, коли у Резолюції A/RES/70/1 «Перетворення нашого світу: Порядок денний сталого розвитку до 2030 року» [16] вперше з'являються опосередковані згадки про цифрові технології в контексті, релевантному для формування підходів до регулювання ШІ. До аналізованого корпусу включено акти, що відповідають принаймні одному з таких критеріїв: (а) безпосередньо згадується штучний інтелект, автоматизоване прийняття рішень, алгоритмічні системи або автономні системи озброєнь як об'єкт регулювання; (б) встановлюються нормативні положення, безпосередньо застосовні до розробки, впровадження або використання систем ШІ, навіть якщо сам термін «штучний інтелект» у ній прямо не вживається – зокрема, положення щодо автоматизованої обробки персональних даних, управління цифровими технологіями або подолання технологічного розриву. При проведенні дослідження не бралися до уваги акти, що стосуються виключно загальних питань інформаційної безпеки, електронної комерції або телекомунікацій без регуляторного зв'язку зі сферою ШІ. На підставі цих критеріїв було охоплено корпус із 23 резолюцій ГА ООН, прийнятих в період 2015–2025 років.

Розподіл резолюцій, присвячених ШІ, за роками їх прийняття є вкрай нерівномірним: у 2015 та 2016 році було прийнято лише по одній резолюції, що опосередковано стосуються тематики регулювання ШІ. Після періоду затишшя, у 2023 році відбувся різкий стрибок до п'яти резолюцій, у 2024 році було прийнято шість резолюцій, а 2025 рік став рекордним: з'явилося десять нових резолюцій, що відносяться до тематики регулювання ШІ. Тобто 70% усіх досліджуваних резолюцій (16 з 23) були прийняті протягом 2024-2025 років, що свідчить про інтенсифікацію міжнародних зусиль у цій сфері (див. Діаграма 1).

Діаграма 1. Динаміка прийняття резолюцій ГА ООН у сфері регулювання ШІ (по роках)



Діаграма 1 наочно демонструє різке зростання нормотворчої активності ГА ООН у сфері регулювання ШІ. Період 2017-2022 років, хоча й не характеризувався прийняттям резолюцій безпосередньо про ШІ, не був простим «провалом» у регуляторній активності. Натомість це був час накопичення практичного досвіду, аналізу викликів і формування міжнародного консенсусу щодо необхідності регулювання зазначеного феномену та, зокрема, роботи на рівні спеціалізованих установ ООН. Різкий стрибок у 2023 році був безпосередньо пов'язаний з масовим впровадженням генеративного ШІ, що актуалізував потребу у правовому регулюванні.

Аналіз змісту прийнятих резолюцій дає підстави виокремити чотири періоди, кожен з яких характеризується специфічними пріоритетами у регулюванні ШІ на рівні ГА ООН.

Перший період (2015-2022): «Підготовча фаза». Цей період характеризується опосередкованими згадками ШІ в контексті загальних цифрових технологій. Основний фокус був зосереджений на широких темах, таких як інформаційно-

комунікаційні технології, подолання цифрового розриву та розбудова інформаційного суспільства. Термін «штучний інтелект» фактично не вживався та не отримав спеціальної концептуалізації. Ключовими резолюціями цього періоду стали Резолюція A/RES/70/1 від 25 вересня 2015 року «Перетворення нашого світу: Порядок денний сталого розвитку до 2030 року» [16] та Резолюція A/RES/70/125 від 1 лютого 2016 року «Підсумковий документ засідання високого рівня Генеральної Асамблеї щодо загального огляду виконання підсумків Всесвітнього саміту з питань інформаційного суспільства» [82]. Відсутність прямих згадок ШІ в Порядку денному до 2030 року, пояснюється тим, що на момент прийняття цієї резолюції системи машинного навчання та нейронні мережі лише починали демонструвати свій потенціал, а їхній вплив на суспільство та економіку ще не був очевидним; водночас сам Порядок денний був сфокусований на загальних аспектах сталого розвитку, в якому конкретні технології розглядалися як інструменти досягнення цілей, а не як самостійні об'єкти регулювання.

Другий період (2023): «Пробудження інтересу», який випав на 2023 рік став переломним у формуванні ставлення міжнародної спільноти до необхідності регулювання ШІ. Цей період характеризується появою перших резолюцій, спеціально присвячених ШІ. Каталізатором таких цих змін стало масове впровадження генеративного ШІ, зокрема систем типу ChatGPT та інших великих мовних моделей (LLM). В цей період ГА ООН були прийняті п'ять резолюцій: A/RES/77/320 «Вплив стрімкого технологічного прогресу на досягнення Цілей сталого розвитку» [18], A/RES/78/241 «Летальні автономні системи озброєнь» [85], A/RES/78/132 «Інформаційно-комунікаційні технології для сталого розвитку» [83], A/RES/78/160 «Наука, технології та інновації для сталого розвитку» [84] та A/RES/77/211 «Право на приватність у цифрову епоху» [17]. В названих резолюціях вперше було зосереджено увагу на етичних аспектах застосування ШІ, питаннях його воєнного застосування та проблемах можливої алгоритмічної дискримінації систем ШІ.

Третій період (2024): «Інституціоналізація». Цей період ознаменувався інтенсифікацією нормотворчої діяльності (загалом було прийнято шість ключових резолюцій) – та якісним переходом від декларування загальних принципів регулювання ШІ до формування спеціалізованої інституційної архітектури для регулювання ШІ. Резолюція A/RES/78/265 «Використання можливостей безпечних, захищених і надійних систем штучного інтелекту для сталого розвитку» [19] запровадила концепцію «життєвого циклу систем штучного інтелекту», сформулювала вимоги щодо безпеки, захищеності та надійності ~~та~~ таких систем і наголосила на необхідності забезпечення прав людини на всіх етапах розробки та використання систем ШІ. Резолюція A/RES/78/311 «Посилення міжнародного співробітництва з нарощування потенціалу в галузі штучного інтелекту» [86] конкретизувала механізми міжнародної співпраці та закріпила заяву «про рішучість подолати розриви у сфері штучного інтелекту» [86, п. 1]. Визначальною ж подією цього періоду стало прийняття Резолюції A/RES/79/1 «Пакт заради майбутнього» [21], додатком до якої є Глобальний цифровий договір (далі – ГЦД) – перший комплексний міжнародно-правовий акт щодо регулювання ШІ на універсальному рівні. ГЦД визначає п'ять стратегічних цілей, з яких п'ята, «посилення міжнародного управління штучним інтелектом на благо людства» [21, п. 7(5)], безпосередньо стосується предмета цього дослідження, тоді як проблематика ШІ наскрізно пронизує й інші цілі Договору: від подолання цифрового розриву та розширення інклюзивності цифрової економіки до просування відповідального управління даними.

До важливих положень ГЦД можна віднести: необхідність «збалансованого, інклюзивного та ризикоорієнтованого підходу до управління новітніми технологіями, із повним та рівним представництвом усіх країн, зокрема країн, що розвиваються, та змістовною участю всіх зацікавлених сторін» [21, п. 48], підкреслює, що «міжнародне співробітництво є необхідним для сприяння координації та забезпечення сумісності систем управління ШІ, що формуються» [21, п. 51], та фіксує стурбованість тим, що «активізація процесів збору даних, обміну

ними та їх обробки, зокрема в системах штучного інтелекту, може призвести до посилення ризиків за відсутності ефективних норм захисту персональних даних та конфіденційності» [21, п. 37]. Водночас Договір наголошує і на позитивному потенціалі ШІ, визнаючи відкриті моделі штучного інтелекту складовою «цифрових суспільних благ» [21, п. 14]. Інституційний вимір ГЦД є найбільш інноваційним його компонентом: він передбачив створення трьох спеціалізованих інституцій – Незалежної міжнародної наукової панелі з ШІ для науково обґрунтованого підходу до регулювання, Глобального діалогу з управління ШІ для міжнародної координації політики та Глобальної рамки управління даними. Він також закріпив підхід за участю багатьох зацікавлених сторін: «ми як уряди працюватимемо у співпраці та партнерстві з приватним сектором, громадянським суспільством, міжнародними організаціями, технічними та науковими колами та всіма іншими зацікавленими сторонами в рамках їх відповідних функцій та обов'язків» [21, п. 6]. Окремою інституційною новацією стало створення Офісу ООН з нових та цифрових технологій (ODET), що замінив попередній Офіс Посланника Генерального секретаря з технологій [21, п. 72].

За юридичною природою ГЦД не є міжнародним договором у розумінні Віденської конвенції про право міжнародних договорів 1969 року, а є додатком до резолюції Генеральної Асамблеї ООН, тобто є «м'якого права». Його юридична сила визначається консенсусним прийняттям 193 державами-членами ООН, що надає йому найвищий рівень політичної легітимності. Водночас сфера дії ГЦД обмежена виключно сферою мирних відносин між державами, оскільки, згідно п. 4 цей Договір «визначає цілі, принципи, зобов'язання та дії, які ми беремо на себе для її досягнення у невоєнній сфері» [21, п. 4]. Отже, критично важлива сфера застосування ШІ у воєнних цілях залишається поза предметом його регулювання.

Незважаючи на консенсусний характер прийняття, переговорний процес щодо ГЦД виявив суттєві розбіжності між окремими державами-членами. Росія, Іран, КНДР та Сирія намагалися доповнити текст положеннями про неутручання зовнішніх суб'єктів у питання національного суверенітету та надання пріоритету

міжурядовим переговорам, вбачаючи у міжнародно-правовому регулюванні загрозу своїй свободі дій щодо розробки та використання технологій ШІ у чутливих сферах державного нагляду й контролю інформаційного простору. М. Наторські констатує, що «існує послідовна опозиція з боку росії, Ірану, КНДР та Білорусі, серед інших, до прийняття різних резолюцій ГА ООН, пов'язаних зі штучним інтелектом» [100, с. 18]. Аналогічна позиція висловлювалася ними під час переговорів щодо Рекомендації ЮНЕСКО, що свідчить про існування стійкої групи держав, які принципово заперечують універсальний підхід до регулювання ШІ. Натомість Група 77 та Китай зайняли інший підхід до цього питання: за свідченням того ж дослідника, ці держави «підкреслили, що Глобальний цифровий договір має бути процесом, керованим державами, та забезпечувати повну та рівну участь усіх держав, особливо країн, що розвиваються», наголошуючи водночас на потребі «зацікавлених сторін дотримуватися національних законів, регуляцій, принципів, правил та норм відповідальної поведінки в ІКТ-середовищі» [100, с. 20]. Таким чином, переговорний процес виявив три відмінні позиції: пряму опозицію рф, Ірану та КНДР, мотивовану захистом національного суверенітету у чутливих сферах; позицію Групи 77 і Китаю, які підтримують регулювання за умови його міждержавного характеру та забезпечення рівного представництва країн, що розвиваються; та позицію, втілену у фінальному тексті ГЦД, яка поєднує міждержавну основу із залученням різних за їх правовим статусом зацікавлених сторін. Ця тріада позицій є характерною для нормотворчих процесів у сфері ШІ та, як буде показано далі, проявляється на інших рівнях регулювання.

Завершення 2024 року ознаменувалося прийняттям трьох додаткових резолюцій. Резолюція A/RES/79/62 «Летальні автономні системи озброєнь» [87] розвинула положення Резолюції A/RES/78/241 щодо збереження людського контролю над рішеннями про застосування летальної сили, а Резолюція A/RES/79/194 «Інформаційно-комунікаційні технології для сталого розвитку» [88] інтегрувала положення про ШІ у ширший контекст цифрової трансформації та досягнення Цілей сталого розвитку. Принципово новим кроком стала Резолюція

A/RES/79/239 «Штучний інтелект у війсьній сфері та його наслідки для міжнародного миру та безпеки» [89] – перша резолюція, що комплексно розглянула використання систем ШІ у війсьних цілях поза контекстом летальних автономних систем озброєнь.

Четвертий період (2025-теперішній час): «Оперативна фаза». 2025 рік став найактивнішим за весь досліджуваний період, оскільки було прийнято десять резолюцій, що становить понад 40% від загальної кількості резолюцій ГА ООН присвячених правовому регулюванню ШІ протягом 2015–2025 років. Характерною рисою цього періоду є операціоналізація інституційної структури, закріпленої ГЦД у 2024 році, через визначення конкретних мандатів, структури та процедур функціонування створюваних органів, а також значне посилення уваги до питань застосування ШІ у війсьній сфері. Початок фази поклала Резолюція A/RES/79/322 «Роль штучного інтелекту у створенні нових можливостей для сталого розвитку в Центральній Азії» [90] – перша регіонально-орієнтована резолюція ГА ООН щодо ШІ. Резолюція A/RES/79/325 «Повноваження та процедури створення та функціонування Незалежної міжнародної наукової панелі з штучного інтелекту та Глобального діалогу з управління штучним інтелектом» [91] стала важливим кроком у реалізації положень ГЦД, встановивши конкретні мандати, структуру та процедури функціонування двох інституційних механізмів, передбачених Договором, а Резолюція A/RES/79/334 «Вплив стрімких технологічних змін на досягнення Цілей та завдань сталого розвитку» [92] розширила аналіз взаємозв'язку між розвитком технологій ШІ та прогресом у досягненні Порядку денного сталого розвитку до 2030 року, наголосивши на ризиках поглиблення нерівності.

Грудень 2025 року виявився особливо насиченим періодом нормотворчої діяльності ГА ООН в сфері регулювання ШІ: з семи прийнятих резолюцій три були присвячені проблематиці застосування ШІ у війсьній сфері. Так, Резолюція A/RES/80/23 «Можливі ризики інтеграції штучного інтелекту в системи командування, управління та зв'язку ядерної зброї» [93] стала першим міжнародно-правовим актом на рівні ГА ООН, що звернув увагу на ризики втрати людського

контролю над системами ядерного командування та можливість неконтрольованої ескалації внаслідок помилок або непередбачуваної поведінки систем ШІ. Резолюція A/RES/80/57 «Летальні автономні системи озброєнь» [94] стала третьою резолюцією ГА ООН з цієї проблематики (після A/RES/78/241 та A/RES/79/62), продемонструвавши послідовне посилення міжнародної уваги до збереження значущого людського контролю над рішеннями про застосування летальної сили, а Резолюція A/RES/80/58 «Штучний інтелект у військовій сфері та його наслідки для міжнародного миру та безпеки» [95] охопила весь спектр застосування систем ШІ у військових цілях – системи розвідки, кібероперації, логістичну підтримку та прийняття рішень на полі бою – та підтвердила застосовність міжнародного гуманітарного права до всіх систем озброєнь на основі ШІ.

Решта грудневих резолюцій інтегрували регулювання ШІ у ширші тематичні контексти інформаційного суспільства та сталого розвитку: Резолюція A/RES/80/84 A-B «Інформація на службі людства» [96] – у контекст доступу до інформації за умов розвитку нових технологій; Резолюції A/RES/80/118 «Інформаційно-комунікаційні технології для сталого розвитку» [97] та A/RES/80/147 «Наука, технології та інновації для сталого розвитку» [98] – у контекст цифрової трансформації та досягнення ЦСР з акцентом на інклюзивний доступ до технологій ШІ для всіх країн та визнанням ШІ ключовим елементом наукового та технологічного прогресу. Завершенням року стало прийняття Резолюції A/RES/80/173 «Підсумковий документ Зустрічі високого рівня Генеральної Асамблеї щодо загального огляду виконання рішень Всесвітньої зустрічі на вищому рівні з питань інформаційного суспільства» (ВВУЮ+20) [99], що підбила підсумки двадцятирічного розвитку глобального інформаційного суспільства та визначила правове регулювання ШІ як центральний елемент цифрової трансформації.

Таким чином, цей період засвідчив остаточне формування ШІ як самостійного предмета міжнародно-правового регулювання на універсальному рівні через операціоналізацію інституційної архітектури ГЦД та формування спеціалізованого режиму для воєнної сфери

Виділені періоди відобразилися не лише в зміні змісту документів, але й у глибокій трансформації термінологічного апарату та риторичних стратегій. У ранніх резолюціях 2015–2022 років домінували широкі узагальнені поняття – «інформаційно-комунікаційні технології», «цифрові технології», «нові та новітні технології», – що відображало загальний підхід до цифрової трансформації, в якому ШІ розглядався як один з елементів ширшого технологічного ландшафту без виокремлення його специфічних правових викликів. Натомість резолюції періоду 2023–2025 років демонструють якісно інший рівень термінологічної точності: з'явилися та закріпилися поняття «системи штучного інтелекту», «безпечні, захищені та надійні системи ШІ», «життєвий цикл систем ШІ», «управління штучним інтелектом», а також спеціалізовані терміни, що фіксують конкретні правові виклики, – «алгоритмічна дискримінація та упередженість», «нагляд з боку людини», «автоматизоване прийняття рішень». Характерною є й трансформація поняття «технологічні розриви» в більш конкретне «розриви в галузі ШІ», що підкреслює специфіку проблеми нерівного доступу саме до технологій ШІ.

Паралельно з термінологічною еволюцією відбувалася трансформація риторичних стратегій. Резолюції 2015–2016 років характеризувалися виразно оптимістичною риторикою, в якій інформаційно-комунікаційні технології «відкривають широкі перспективи» та «сприяють економічному розвитку», без достатньої концептуалізації можливих ризиків. Резолюції 2023 року ознаменували перехід до збалансованої риторики, яка водночас визнавала як «потенційні переваги», так і «серйозні проблеми та питання, що викликають занепокоєння», відображаючи зростаюче усвідомлення двоїстої природи технологій ШІ та необхідності одночасного стимулювання інновацій і управління ризиками. Резолюції 2024–2025 років демонструють обережно-проактивну риторичну поєднує визнання «нагальної необхідності» регулювання з наголосом на «критичних ризиках» та одночасно підтримує оптимістичне бачення використання ШІ «на благо всіх»; особливо показовою є конструкція «безпечні, захищені та надійні системи штучного інтелекту», яка одночасно визнає існування небезпечних систем і

встановлює нормативний ідеал, до якого має прагнути міжнародна спільнота. Найбільш виразною є еволюція риторики щодо воєнного застосування ІІІ: від обережних згадок про «потенційні ризики» у 2023 році до категоричних формулювань про «критичні загрози», «екзистенційні ризики» та «неприйнятні наслідки» втрати людського контролю над автономними системами озброєнь у 2024–2025 роках. Водночас риторика останніх резолюцій послідовно акцентує «інклюзивність» та «багатосторонність», наголошуючи на «повній та рівній участі всіх держав, особливо країн, що розвиваються», що відображає намагання міжнародної спільноти протистояти ризикам монополізації розробки ІІІ обмеженою кількістю технологічно розвинених держав та корпорацій.

Статистичний аналіз частоти ключових термінів підтверджує цю еволюцію: «сталий розвиток» залишається найчастотнішим з 154 згадками в 88% досліджуваних резолюцій, що свідчить про послідовне позиціонування ІІІ як інструменту досягнення Цілей сталого розвитку. Високу частотність демонструють терміни «міжнародна співпраця» (89 згадок), «безпека» (78 згадок) та «нарощування потенціалу» (73 згадки). Водночас терміни, пов'язані з етичними та правовими викликами, такі як «права людини» (67 згадок), «приватність» (52 згадки), «прозорість» (45 згадок) та «дискримінація» (29 згадок), також набули значної присутності, особливо в резолюціях останніх років. Детальний контент-аналіз прийнятих ГА ООН резолюцій дозволяє виділити десять основних тематичних категорій, що відображають пріоритети міжнародної спільноти у сфері регулювання ІІІ. Кожна категорія представлена в різній кількості резолюцій, що свідчить про ступінь уваги до відповідних питань.

Таблиця 1. Тематична класифікація резолюцій ГА ООН щодо ІІІ

№	Тематична категорія	Кількість резолюцій	Ключові принципи
1	Міжнародна співпраця	20	Багатостороннє співробітництво, передача технологій

2	Етика та права людини	18	Людиноцентричний підхід, недискримінація
3	Подолання цифрового розриву	17	Інклюзивність, рівний доступ
4	Прозорість і підзвітність	16	Пояснюваність, алгоритмічна справедливість
5	Сталий розвиток	15	Зв'язок з ЦСР, екологічна відповідальність
6	Застосування у воєнній сфері	12	Людський контроль, заборона автономних систем III
7	Безпека та надійність	11	Технічна безпека, управління ризиками
8	Приватність і захист даних	10	Конфіденційність, згода на обробку
9	Економічний розвиток	9	Інновації, конкурентоспроможність
10	Нарощування потенціалу	8	Освіта, професійна підготовка

Тематична класифікація демонструє намагання охопити як технічні аспекти (безпека, надійність), так і соціально-етичні виміри (права людини, етика), економічні питання (інновації, розвиток) та геополітичні проблеми (міжнародна співпраця, цифровий розрив). Домінування теми міжнародної співпраці, представлена у 20 з 23 резолюцій, відображає глибоке розуміння того, що жодна держава не здатна самостійно забезпечити ефективне регулювання транскордонної за природою технології. Високі позиції тем етики та прав людини (18 резолюцій), подолання цифрового розриву (17 резолюцій), прозорості та підзвітності (16 резолюцій) і сталого розвитку (15 резолюцій) свідчать про стурбованість міжнародної спільноти захистом основних прав, забезпеченням рівного доступу до технологій III між розвиненими країнами та країнами, що розвиваються, та зв'язком регулювання III з Цілями сталого розвитку. Окремий тематичний блок утворюють

резолюції з воєнного застосування ШІ (12 резолюцій), а технічні та інфраструктурні питання – безпека та надійність (11), приватність та захист даних (10), економічний розвиток (9) і нарощування потенціалу (8) – представлені у меншому, але стабільному обсязі.

Проведений аналіз резолюцій ГА ООН дозволяє виокремити дванадцять ключових принципів, що становлять основу міжнародно-правового регулювання ШІ на універсальному рівні. Ці принципи не лише систематично повторюються в різних резолюціях, але й демонструють поступову консолідацію та деталізацію, що свідчить про формування сталого консенсусу щодо базових вимог до розробки, впровадження та використання систем ШІ.

Принцип 1: Людиноцентричний підхід. Усі 23 досліджувані резолюції послідовно наголошують на необхідності поважати, захищати та заохочувати права людини та основні свободи протягом усього життєвого циклу систем ШІ. Резолюція A/RES/78/265 «підкреслює, що права людини та основні свободи повинні поважатися, захищатися та заохочуватися протягом усього життєвого циклу систем штучного інтелекту» [19, п. 5], а ГЦД відтворює це положення на рівні універсального консенсусу, зобов'язуючи держави «забезпечити дотримання норм міжнародного права прав людини протягом усього життєвого циклу цифрових та новітніх технологій» [21, п. 22].

Принцип 2: Безпека, захищеність та надійність. Резолюції наголошують на необхідності створення систем ШІ безпечними, захищеними та надійними протягом усього їхнього життєвого циклу. Резолюція A/RES/78/265 запровадила базове визначення, згідно з яким «безпечні, захищені та надійні системи ШІ – під якими для цілей цієї резолюції розуміються системи ШІ в не воєнній сфері, життєвий цикл яких включає етапи попереднього проектування, проектування, розробки, оцінки, тестування, впровадження, використання, продажу, закупівлі, експлуатації та виведення з експлуатації» [19]. Це визначення було дослівно відтворене у Преамбулі Резолюції A/RES/78/311 [86] та консолідоване у наступних актах, що засвідчує формування стандартизованої термінологічної конструкції на універсальному рівні

з акцентом на технічній безпеці, кібербезпеці та управлінні ризиками високоризикових систем ШІ.

Принцип 3: Прозорість та пояснюваність. Цей принцип передбачає сприяння розробці технічних засобів, що дозволяють користувачам виявляти маніпуляції з інформацією та забезпечують зрозумілість алгоритмічних рішень. ГЦД «закликає компанії, що працюють у сфері цифрових технологій, та платформи соціальних мереж підвищити транспарентність та підзвітність своїх систем, у тому числі умов надання послуг, процедур модерації контенту та принципів роботи алгоритмів рекомендацій та способів обробки особистих даних користувачів на місцевих мовах, щоб дати користувачам можливість приймати усвідомлені рішення та давати або відкликати інформовану згоду» [21, п. 36(a)]. Аналогічну вимогу формулює Резолюція A/RES/78/265, закликаючи до розробки «ефективних та сумісних на міжнародному рівні рамок, методів та стандартів... з метою недопущення посилення або увічнення дискримінаційних чи упереджених видів застосування» [19, п. 4(h)].

Принцип 4: Підзвітність та людський нагляд. Цей принцип передбачає, що розробники та оператори систем ШІ несуть відповідальність за їхнє функціонування та наслідки використання. Резолюція A/RES/80/58 у Преамбулі «підкреслює важливість забезпечення відповідального застосування штучного інтелекту у війсьній сфері, яке для цілей цієї резолюції включає людиноцентричне, підзвітне, безпечне, надійне та гідне довіри використання штучного інтелекту відповідно до норм міжнародного права» [95]. Людський нагляд означає збереження значущого контролю людини над критичними рішеннями, що приймаються з використанням ШІ. Ця ж резолюція визнає «необхідність застосування державами належних гарантій, включаючи заходи, що стосуються прийняття рішень та забезпечення контролю з боку людини в питаннях застосування сили, для забезпечення відповідального застосування штучного інтелекту у війсьній галузі згідно з відповідними зобов'язаннями держав, що впливають з норм міжнародного права, що застосовуються» [95]. Цей принцип, артикульований на універсальному рівні в резолюціях ГА ООН, згодом дістав конкретизацію в юридично обов'язкових актах

на регіональному рівні: як зазначено в попередніх авторських публікаціях, його нормативне закріплення «характеризується переходом від актів м'якого права до юридично обов'язкових документів» [286, с. 235], а якісно новим етапом стало закріплення принципу в Регламенті (ЄС) 2024/1689 у формі концепції «відповідного контролю людини» для високоризикових систем. Самі механізми забезпечення контролю людини над системами ШІ доцільно поділити, як обґрунтовано в попередніх авторських дослідженнях, на превентивні (ex-ante) та реактивні (ex-post): «Міжнародно-правова практика демонструє формування комплексу правових механізмів забезпечення контролю людини над системами ШІ, які можна умовно поділити на превентивні (ex-ante) та реактивні (ex-post)» [60, с. 69]. До перших належать обов'язкові оцінки впливу на права людини, вимоги до проектування систем та їх сертифікація, до других – механізми відповідальності та оскарження. Превентивні, що спрямовані на запобігання порушенням ще на етапах проектування й розробки, систематизовано у чотири групи: «інструменти оцінки впливу на права людини, вимоги до проектування систем, сертифікація та стандартизація, а також документування та реєстрація» [287, с. 150]. Такий підхід узгоджується із загальною тенденцією міжнародного права до пріоритизації запобігання над відшкодуванням.

Принцип 5: Інклюзивність та недискримінація. Цей принцип вимагає забезпечення того, щоб системи ШІ не посилювали існуючі форми дискримінації та не створювали нових. Резолюція A/RES/78/311 «рекомендує державам-членам приймати профілактичні заходи щодо протидії расизму, дискримінації та іншим формам алгоритмічної упередженості» [86, п. 7], а Резолюція A/RES/77/211 поширює відповідні зобов'язання на приватний сектор, закликаючи «всі комерційні підприємства активізувати зусилля з боротьби з дискримінацією, що виникає внаслідок використання систем штучного інтелекту, в тому числі шляхом проявлення належної обачності при оцінці, запобіганні та пом'якшенні несприятливого впливу на права людини при їх застосуванні» [17, п. 8(g)]. Особлива увага приділяється боротьбі з алгоритмічною упередженістю та забезпеченню рівного доступу до переваг технологій ШІ для всіх груп населення.

Принцип 6: Міжнародне співробітництво. Резолюція A/RES/78/311 «заявляє про рішучість ... розширити міжнародне співробітництво в галузі нарощування потенціалу в країнах, що розвиваються» [86, п. 1] та закликає міжнародне співтовариство створювати сприятливі умови для такого співробітництва з «дотриманням норм міжнародного права, зокрема Статуту ООН, а також прав людини та основних свобод» [86, п. 3]. ГЦД конкретизує цю установку, підкреслюючи, що «міжнародне співробітництво необхідне для сприяння координації та забезпечення сумісності систем управління штучним інтелектом, що формуються» [21, п. 51].

Принцип 7: Подолання цифрового розриву. Численні резолюції наголошують на необхідності забезпечення справедливого доступу країн, що розвиваються, до технологій ШІ. Резолюція A/RES/78/311 «із занепокоєнням відзначає, що розриви в галузі ШІ та інших цифрових технологій між країнами та всередині країн продовжують зростати, а країни, що розвиваються, стикаються з особливими проблемами, не встигаючи за стрімким прискоренням темпів розвитку технології ШІ» [86] та закликає розвинені країни «розширювати співробітництво в справі нарощування потенціалу, в тому числі у форматі обміну думками з питань політики, заходів з обміну знаннями та передачі технологій на взаємоузгоджених умовах» [86, п. 5]. Резолюція A/RES/79/322 [90] вперше адаптує цю установку до регіональної специфіки, акцентуючи особливі потреби країн Центральної Азії.

Принцип 8: Сприяння сталому розвитку. ШІ розглядається як потужний інструмент для досягнення Цілей сталого розвитку. Резолюція A/RES/78/265 «визнає, що безпечні, захищені та надійні системи ШІ здатні прискорити та стимулювати прогрес у досягненні всіх 17 цілей у галузі сталого розвитку та забезпеченні сталого розвитку в усіх його трьох – економічному, соціальному та екологічному – компонентах на збалансованій та комплексній основі» [19]. Аналогічну позицію відтворюють Резолюції A/RES/78/311 [86], A/RES/79/334 [92] та A/RES/80/147 [98], послідовно закликаючи використовувати ШІ для прискорення

прогресу в реалізації ЦСР, особливо у сферах охорони здоров'я, освіти, боротьби зі зміною клімату та подолання бідності.

Принцип 9: Захист приватності та даних. Резолюції підкреслюють необхідність захисту права на приватність в епоху ШІ. Резолюція A/RES/77/211 «закликає держави захищати людей від порушення або обмеження права на недоторканність приватного життя, в тому числі внаслідок довільного або неправомірного збору, обробки, зберігання, поширення та структуризації даних, а також використання автоматизованих процесів та машинного навчання» [17, п. 6(о)] та рекомендує систематичне проведення «регулярних та всеосяжних оцінок впливу на права людини та участь усіх відповідних зацікавлених сторін» [17, п. 10] протягом усього життєвого циклу систем ШІ. ГЦД доповнює ці положення застереженням про те, що «активізація процесів збору даних, обміну ними та їх обробки, в тому числі в системах ШІ, може призвести до посилення ризиків за відсутності ефективних норм захисту персональних даних та конфіденційності» [21, п. 37].

Принцип 10: Підхід за участю багатьох зацікавлених сторін. ГЦД закріпив підхід, що передбачає активну участь приватного сектору, технологічних компаній, наукової спільноти та громадянського суспільства у процесі вироблення міжнародних стандартів регулювання ШІ. Резолюція A/RES/78/265 «підкреслює необхідність того, щоб уряди, приватний сектор, громадянське суспільство, міжнародні та регіональні організації, наукові та дослідницькі установи, технічні спільноти та всі інші зацікавлені сторони продовжували працювати разом, відповідно до обставин» [19, п. 8]. Резолюція A/RES/80/58 поширює цей підхід на воєнну сферу, визнаючи внесок усіх категорій зацікавлених сторін, включно з Міжнародним комітетом Червоного Хреста, у формування політики застосування ШІ у воєнних цілях та підкреслюючи «важливість підходу за участю багатьох зацікавлених сторін» [95].

Принцип 11: Обережність (запобіжний підхід). Найвиразніше цей принцип відтворюють резолюції щодо воєнного застосування ШІ, логіка яких відповідає структурі запобіжного підходу: за наявності ризику катастрофічної та необоротної

шкоди невизначеність щодо поведінки систем ШІ не може виправдовувати їх безконтрольне застосування. Резолюція A/RES/80/23 «висловлює занепокоєння тим, що використання штучного інтелекту в процесі прийняття рішень, що стосуються систем командування, управління та зв'язку комплексу ядерної зброї, могло б послабити людський контроль і нагляд, підвищуючи ймовірність виникнення зумовлених спотворень у середовищах прийняття рішень та скорочення строків вжиття заходів і контрзаходів, особливо на таких найбільш чутливих та критично важливих етапах, як прийняття рішення про пуск, що може підвищити ризик випадкового, ненавмисного або несанкціонованого застосування ядерної зброї» [93]. Резолюція додатково звертає увагу на технічні обмеження систем ШІ – «можливість збоїв, експлуатації вразливостей або злому, а також когнітивні відхилення та упередженість автоматики», які можуть породжувати «галюцинації» та призводити до «серйозних та катастрофічних наслідків» [93]. Саме превентивний характер вимоги – резолюція вимагає збереження людського контролю та нагляду ще до настання будь-якої шкоди – відрізняє цей принцип від суміжного принципу людського нагляду.

Принцип 12: Застосовність міжнародного права. Резолюція A/RES/80/58 підтверджує, що «норми міжнародного права, включаючи Статут ООН, міжнародне гуманітарне право та міжнародне право прав людини, застосовуються до регульованих ними питань, які виникають протягом усього життєвого циклу засобів ШІ, а також до систем, функціонування яких у воєнній галузі вони забезпечують» [95, п. 1]. Резолюція A/RES/78/265 «підтверджує міжнародне право, зокрема Статут ООН» та «визнає, що безпечні, захищені та надійні системи ШІ повинні бути антропоцентричними, надійними, пояснюваними, етичними, інклюзивними, повністю забезпечувати дотримання, заохочення та захист прав людини та міжнародного права» [19]. Резолюція A/RES/78/311 додатково наголошує, що системи ШІ «повинні ґрунтуватися на всебічному дотриманні, заохоченні та захисті прав людини та норм міжнародного права» [86].

На підставі проведеного індуктивного узагальнення – систематичного порівняння елементів тексту резолюцій, термінологічних конструкцій та нормативних формулювань у всіх 23 досліджуваних резолюціях – вдалося виявити три крос-актові закономірності. Метод полягав у простеженні того, як конкретна правова конструкція, що з'являється в одному акті, відтворюється в наступних, набуваючи при цьому ознак стандартизованого нормативного елемента. Це не суто статистичний підрахунок частотності термінів, а якісний аналіз міграції нормативних формулювань між резолюціями.

Першою виявленою закономірністю у досліджуваних резолюціях є систематизація концепту «життєвого циклу систем штучного інтелекту» як базової регуляторної рамки. Цей концепт, вперше чітко артикульований у Резолюції A/RES/78/265 у березні 2024 року, передбачає, що «права людини та основні свободи мають поважатися, захищатися та заохочуватися протягом усього життєвого циклу систем штучного інтелекту» [19, п. 5]. Примітно, що він не просто згаданий у подальших актах, а систематично інкорпорований спочатку у ГЦД (A/RES/79/1) [21], а потім послідовно відтворений у резолюціях A/RES/79/194 [88], A/RES/80/58 [95] та A/RES/80/118 [97]. Така швидка консолідація свідчить про формування єдиної правової рамки, що розглядає регулювання ШІ не як одномоментне втручання, а як безперервний процес управління ризиками протягом усього життєвого циклу технології.

Другою закономірністю є стандартизація тріади «безпечні, захищені та надійні системи штучного інтелекту» (англ. – *safe, secure and trustworthy AI systems*). Ця формула, також вперше чітко артикульована в Резолюції A/RES/78/265 [19], стала стандартним способом позначення бажаних характеристик систем ШІ. Кожен з трьох елементів тріади несе специфічне нормативне навантаження: «безпечні» стосується технічної надійності та відсутності ризиків для користувачів, «захищені» – кібербезпеки та стійкості до зловмисних втручань, «надійні» – передбачуваності функціонування та відповідності очікуванням користувачів. Систематичне повторення цієї тріади в резолюціях A/RES/79/1 [21], A/RES/78/311 [86],

A/RES/79/334 [92] та інших свідчить про формування усталеної термінологічної конструкції, яка може з часом набути характеру міжнародно-правового стандарту.

Третьою закономірністю є позиціонування ГЦД як точки конвергенції міжнародно-правового регулювання ШІ на універсальному рівні: після прийняття Резолюції A/RES/79/1 [21] усі наступні резолюції без винятку посиляються на ГЦД, узгоджують свої положення з його принципами та розвивають окремі аспекти встановленої ним регуляторної рамки. Це особливо помітно в резолюціях про застосування ШІ у війсьній сфері – A/RES/80/23 [93] та A/RES/80/58 [95], які конкретизують положення ГЦД щодо необхідності збереження контролю людини в контексті систем озброєнь, та в Резолюції A/RES/79/325 [91] про нарощування потенціалу, що деталізує передбачені ГЦД механізми міжнародної співпраці.

Ці закономірності в сукупності свідчать про перехід від фрагментарного до системного регулювання ШІ. Міжнародна спільнота рухається від простого накопичення окремих міжнародно-правових актів до формування когерентної нормативної системи, в якій різні елементи підтримують та взаємодоповнюють один одного. Процес консолідації правових позицій відбувається не через формальну уніфікацію термінології, а через органічне закріплення концептуальних рамок, що демонструють свою практичну придатність та здобувають визнання широкого кола держав-членів ООН.

Попри значний прогрес у розвитку міжнародно-правового регулювання ШІ на цьому рівні, аналіз змісту резолюцій виявляє низку суттєвих прогалин. Передусім слід вказати на недостатнє врегулювання ролі приватного сектору, зокрема транснаціональних технологічних корпорацій: резолюції обмежуються загальними посиланнями на Керівні принципи ООН щодо бізнесу та прав людини та декларативними закликами до корпоративної відповідальності, не встановлюючи конкретних механізмів підзвітності розробників систем ШІ високого ризику чи протидії монополізації ринку. М. Наторскі констатує: «незважаючи на сотні керівних принципів, рамкових документів та принципів, прийнятих урядами, компаніями, консорціумами, регіональними та міжнародними організаціями,

відсутня всеохоплююча та репрезентативна глобальна система координації та впровадження глобальної політики у сфері ШІ» [100, с. 3]. Тісно пов'язана з цим прогалина у регулюванні двох найдинамічніших сегментів сучасного ШІ – генеративного ШІ та біометричних технологій: резолюції не розглядають питання захисту авторських прав на згенерований контент, використання захищених матеріалів для навчання моделей, поширення дипфейків чи проблеми «галюцинацій», а єдиним опосередкованим положенням є згадка в Резолюції A/RES/78/265 про «заохочення розробки та впровадження ефективних, доступних, адаптованих, сумісних на міжнародному рівні технічних засобів, стандартів чи методів... які дозволяють користувачам виявляти маніпуляції з інформацією» [19, п. 4(g)], проте воно сформульоване надто загально. У сфері біометричних технологій резолюції не встановлюють специфічних обмежень на масове біометричне спостереження, розпізнавання облич у громадських місцях чи біометричну категоризацію за расовими, етнічними чи гендерними ознаками; Резолюція A/RES/77/211 [17] містить лише загальне застереження про необхідність захисту приватності, що залишає простір для зловживань у контексті масового нагляду.

Друга прогалина – недостатнє регулювання екологічного впливу технологій ШІ та їхнього впливу на трудові відносини. Резолюції A/RES/80/147 [98] та A/RES/79/194 [88] містять лише епізодичні згадки про необхідність врахування екологічних викликів розвитку ШІ, тоді як жодна резолюція не встановлює вимог щодо зменшення викидів парникових газів, обов'язкової звітності про вуглецевий слід технологій або створення стимулів для розвитку «зеленого ШІ». Між тим масштаб проблеми засвідчений емпіричними даними: навчання однієї системи ШІ генерує викиди парникових газів, еквівалентні «викидам п'яти автомобілів протягом усього терміну їх експлуатації» [102]; дата-центри, що обслуговують системи ШІ, споживають близько 2% світового виробництва електроенергії [103]; виробництво спеціалізованих чіпів потребує значних обсягів ультрачистої води [104] та породжує суттєві викиди CO₂-еквіваленту [105]. Показово, що декларативність екологічного виміру притаманна не лише універсальному рівню: аналізуючи Акт про ШІ, М. О.

Медведева доходить висновку, що «екологічна стійкість наразі залишається ... переважно декларативною: відсутні прямі зобов'язання враховувати кліматичні ризики при класифікації систем як високоризикових, не запроваджено обов'язкової «кліматичної» звітності щодо ШІ, а критерії енерго- та ресурсоефективності не інтегровані в процедури оцінки відповідності» [288, с. 386]. Те, що навіть найбільш розроблений юридично обов'язковий акт відтворює цю прогалину, підтверджує загальнішу закономірність, виявлену в нашому дослідженні: жоден із досліджуваних рівнів регулювання не забезпечує адекватного захисту довкілля від впливу ШІ. Подібно недостатньо розробленими залишаються механізми захисту прав працівників в умовах автоматизації: резолюції обмежуються загальними згадками про «зміни в сфері праці» без встановлення конкретних зобов'язань щодо соціального захисту.

Третя прогалина має найбільш фундаментальний характер та стосується довгострокових, екзистенційних та інших системних викликів: усі проаналізовані резолюції фокусуються на вузькому штучному інтелекті, тоді як потенційні ризики створення систем штучного загального інтелекту (AGI) чи суперінтелекту не отримали жодної уваги в документах ГА ООН. Поза увагою резолюцій залишилися й суміжні технологічні виклики – ризики поєднання квантових обчислень та ШІ для криптографічних систем, етичні виклики нейроінтерфейсів та використання ШІ в контексті синтетичної біології. До низки викликів, що залишаються поза увагою міжнародно-правового регулювання на універсальному рівні, належить вплив ШІ на формування світогляду людини – викривлення інформаційного середовища та маніпулятивний вплив на її переконання й поведінку, що зачіпає когнітивну автономію особи. Резолюції ГА ООН порушують це питання лише дотично, обмежуючись питанням виявлення «маніпуляцій з інформацією» [19, п. 4(g)], тоді як предметніша реакція простежується вже на регіональному рівні – у забороні маніпулятивних практик Актом про ШІ (ст. 5) [25] та в актах Ради Європи, зокрема Декларації про маніпулятивні спроможності алгоритмічних процесів [215]. Утім, цілісного міжнародно-правового режиму захисту людини від маніпулятивного

впливу ШІ досі не сформовано на жодному рівні. А. Верчеллі слушно зазначає: «як визнала ООН, штучний інтелект має екзистенційні наслідки. На жаль, позиція ООН щодо ШІ все ще залишається двоякою: вона, здається, коливається між зростаючими гуманітарними потребами світу в полікризі (пандемія, зміна клімату та війни) та очевидним порушенням прав людини, що стоїть за масовим використанням ШІ» [101, с. 4]. Така двоїста позиція, поєднана з відсутністю обов'язкових механізмів імплементації, підкреслює необхідність подальшого вдосконалення міжнародно-правового регулювання ШІ та продовження системної роботи, зокрема на рівні спеціалізованих установ ООН та міжрегіональних організацій.

2.2. Ініціативи спеціалізованих установ ООН у сфері регулювання штучного інтелекту

Аналіз резолюцій Генеральної Асамблеї ООН засвідчив формування загальної нормативної рамки міжнародно-правового регулювання ШІ на універсальному рівні. Резолюції встановили фундаментальні принципи, створили базову термінологію та окреслили напрямки міжнародної співпраці у цій сфері. Водночас паралельно з роботою ГА ООН активну нормотворчу діяльність провадять спеціалізовані установи системи ООН, кожна з яких розробляє специфічні регуляторні підходи у контексті свого функціонального мандату. ЮНЕСКО зосередила зусилля на етичних засадах використання ШІ, Міжнародний союз електров'язку – на технічній стандартизації систем ШІ, Всесвітня організація охорони здоров'я – на галузевих стандартах безпеки медичних систем ШІ, Всесвітня організація інтелектуальної власності досліджує правові наслідки ШІ для патентного та авторського права, а Міжнародна організація праці аналізує соціально-економічний вплив систем ШІ на ринок праці. Ці ініціативи конкретизують загальні принципи, встановлені резолюціями ГА ООН, через секторальні механізми та стандарти, адаптовані до специфіки відповідних сфер міжнародної співпраці.

Перш ніж перейти до аналізу діяльності окремих установ, необхідно з'ясувати юридичну природу актів, що приймаються спеціалізованими установами ООН у сфері ШІ, оскільки саме вона визначає їхню регуляторну дієвість. На відміну від резолюцій ГА ООН, які формально є рекомендаціями загального характеру та адресуються всім державам-членам, акти спеціалізованих установ охоплюють широкий спектр правових форм – від рекомендацій та резолюцій вищих органів установ до технічних стандартів, керівництв та аналітичних документів. Ці акти утворюють своєрідну «градацію нормативності»: рекомендації Генеральної конференції ЮНЕСКО, одностайно прийняті 193 державами-членами, мають найвищий рівень політичного та морального авторитету серед проаналізованих інструментів; резолюції Повноважної конференції МСЕ створюють програмні зобов'язання для органів Союзу та формують мандат діяльності у телекомунікаційному секторі; технічні стандарти (рекомендації) МСЕ, хоча формально добровільні, стають де-факто обов'язковими завдяки вимогам сумісності систем; керівництва ВООЗ адресовані національним регуляторам та медичним установам і мають характер експертних настанов; аналітичні документи ВОІВ виконують функцію картографування правових проблем; а дослідження МОП формують доказову базу для подальшої нормотворчої діяльності. Жоден із цих актів не є міжнародним договором у розумінні Віденської конвенції про право міжнародних договорів 1969 року та не є у вичерпному переліку джерел міжнародного права відповідно до ст. 38(1) Статуту Міжнародного суду ООН, однак у сукупності вони формують щільну мережу «м'якого права», що набуває дедалі більшого регуляторного значення. Як зауважують дослідники Й. Таллберг, Е. Ерман та М. Фурендаль, міжнародно-правове регулювання ШІ «ймовірно буде домінуватися «надм'яким правом», зважаючи на те, що наразі існують численні процеси створення глобальних стандартів поза традиційними форумами міжнародного правотворення» [146, с. 8]. Як зазначають Г. Робертс, М. Таддео та Л. Флоріді, відмінність між «м'яким» і «твердим правом» при цьому слід розглядати «як континуум за трьома вимірами – ступенем зобов'язання, точності та делегування

– а не як бінарну дихотомію» [150, с. 8]. Такий підхід дозволяє оцінити реальний нормативний потенціал кожного акту спеціалізованих установ незалежно від його формальної назви – рекомендація, керівництво, технічний стандарт чи аналітичний звіт.

Таблиця 2. Функціональні напрями роботи спеціалізованих установ ООН у сфері регулювання ШІ

Установа	Мандат у сфері ШІ	Тип ініціативи / міжнародно-правового акту	Охоплення
ЮНЕСКО	Етичні засади, цінності, культурне різноманіття, освіта	Рекомендації, методології, оцінки	194 держави-члени та 12 асоційованих членів
МСЕ	Технічна стандартизація ІКТ, впровадження для ЦСР	Технічні стандарти, резолюції, фокус-групи	194 держави-члени та понад 1000 організацій
ВООЗ	Охорона здоров'я, медичні системи ШІ	Керівництва, рекомендації, регуляторні вказівки	194 держави-члени
ВОІВ	Інтелектуальна власність, патенти, авторське право	Діалоги, аналітичні звіти, бази даних політик	194 держави-члени
МОП	Праця, зайнятість, соціальний захист	Дослідження, аналітичні звіти, обсерваторія	187 держав-членів

Хронологічний аналіз ініціатив свідчить про послідовне формування цілісної регуляторної екосистеми, започаткованої у 2017 році разом із запуском під егідою МСЕ платформи «Глобальний Саміт ШІ заради добра» (англ. – AI for Good Global Summit). У 2019 році ЮНЕСКО отримала мандат ГА ООН на розробку

універсального нормативного акту з етики ШІ, а ВОІВ ініціювала Діалог з питань інтелектуальної власності в контексті технологій ШІ. У листопаді 2021 року ЮНЕСКО прийняла Рекомендацію про етику штучного інтелекту [13], схвалену всіма 193 державами-членами, одночасно з випуском ВООЗ перших керівництв з етики та регулювання ШІ у сфері охорони здоров'я. Розглянемо кожну ініціативу більш детально.

Організація Об'єднаних Націй з питань освіти, науки і культури (ЮНЕСКО) розробила Рекомендацію про етику штучного інтелекту [13], що посідає центральне місце в системі актів спеціалізованих установ ООН щодо регулювання ШІ. З позиції міжнародного публічного права цей акт є рекомендацією Генеральної конференції ЮНЕСКО, ухваленою відповідно до статті IV(4) Статуту Організації. На відміну від конвенцій ЮНЕСКО, рекомендації не потребують ратифікації та не створюють юридично обов'язкових зобов'язань для держав, маючи характер «бажаних стандартів поведінки», що визначає як переваги (універсальність підтримки, гнучкість імплементації), так і обмеження (відсутність механізмів примусового виконання) Рекомендації. Її можна кваліфікувати як джерело «м'якого права» з високим ступенем легітимності, зумовленим консенсусним характером прийняття та універсальним членством організації. Як наголошують Й. Таллберг, Е. Ерман та М. Фурендаль, «хоча такі м'які документи можуть спочатку бути розроблені як неправові тексти, на практиці вони можуть набути значної сили у структуруванні міжнародних відносин» [146, с. 8].

За змістом Рекомендація ґрунтується на чотирьох фундаментальних цінностях та десяти керівних принципах. Першою цінністю проголошено повагу, захист та заохочення прав людини та людської гідності [13, п. 13], причому Рекомендація встановлює категоричний імператив: «жодна людина чи людська спільнота не повинні зазнавати шкоди чи підпорядкування, будь то фізичного, економічного, соціального, політичного, культурного чи психічного, на будь-якій фазі життєвого циклу систем ШІ» [13, п. 14]. Друга цінність – процвітання навколишнього середовища та екосистем – зобов'язує всіх суб'єктів життєвого циклу систем ШІ

дотримуватися відповідного міжнародного та національного законодавства щодо захисту навколишнього середовища [13, п. 18]. Третя цінність – забезпечення різноманітності та інклюзивності протягом життєвого циклу систем ШІ «відповідно до міжнародного права, включаючи право прав людини» [13, п. 19]. Четверта цінність – життя в мирних, справедливих та взаємопов'язаних суспільствах – наголошує на партисипативній ролі суб'єктів ШІ у забезпеченні «мирних та справедливих суспільств... на благо всіх, відповідно до прав людини та основоположних свобод» [13, п. 22].

Серед десяти керівних принципів особливе значення для міжнародно-правового регулювання ШІ мають принципи, безпосередньо пов'язані з обмеженням автономії систем ШІ та захистом прав людини. Як підкреслює Т. Біркштедт, «дослідники, здається, досягли консенсусу, що існуючі етичні принципи повинні мати на меті артикулювати загальнолюдські цінності» [111, с. 148]. Принцип пропорційності та незавдання шкоди вимагає, щоб процеси життєвого циклу систем ШІ «не перевищували те, що необхідно для досягнення легітимних цілей або завдань» [13, п. 25], та формулює заборону використання ШІ «для соціального скорингу або масового нагляду» [13, п. 26]. Принцип людського нагляду та визначення вимагає можливості «віднести етичну та правову відповідальність за будь-який етап життєвого циклу систем ШІ до фізичних осіб або до існуючих юридичних осіб» [13, п. 35], і категорично наголошує: «рішення про життя і смерть не повинні передаватися системам ШІ» [13, п. 36]. Принцип прозорості та пояснюваності зобов'язує інформувати людей, «коли рішення приймається на основі алгоритмів ШІ, особливо коли це впливає на їхню безпеку або права людини» [13, п. 38], а принцип відповідальності та підзвітності передбачає розробку «відповідних механізмів нагляду, оцінки впливу, аудиту та належної обачності, включаючи захист інформаторів» [13, п. 43]. Додатково Рекомендація закріплює принципи безпеки та захищеності [13, п. 27], справедливості та недискримінації [13, п. 28–29], сталого розвитку [13, п. 31], права на приватність та захист даних [13, п. 32–33], обізнаності

та грамотності [13, п. 44], а також підходу за участю багатьох зацікавлених сторін та адаптивного управління [13, п. 47].

Водночас сам факт нормативного закріплення принципів не означає їхню автоматичну операціоналізацію. Концепція «принципізму» – використання етичних принципів як нормативного базису – стала, за оцінкою Н. Корреа, «резонуючою основою для більшості різноманітних етичних рекомендацій, політичних рамок, правових орієнтирів, кодексів поведінки, практичних посібників, тощо» [134, с. 4], однак механізми переведення абстрактних принципів у конкретні правові норми залишаються нерозробленими. Дж. Крайгер підтверджує цю проблему, наголошуючи на «чіткій потребі в дослідженнях, які сприяють переходу від того «що» являє собою етика ШІ до того «як» здійснювати управління та операціоналізацію, оскільки наявність добрих етичних принципів... не є достатньою для забезпечення етичної розробки та впровадження штучного інтелекту» [135, с. 355]. Усвідомлюючи цю прогалину, ЮНЕСКО розробила два інструменти імплементації. Методологія оцінки готовності (англ. – Readiness Assessment Methodology, RAM) [106] пропонує державам-членам систему з понад 140 індикаторів у п'яти вимірах – правовому, соціокультурному, економічному, науково-освітньому та технологічному – для оцінки національної спроможності впроваджувати етичні принципи регулювання ШІ; станом на кінець 2025 року цю методологію застосовують у понад 80 країнах світу [133]. Методологія оцінки етичного впливу (англ. – Ethical Impact Assessment, EIA) [107] дозволяє заздалегідь виявляти потенційні етичні ризики ще на етапі проектування систем ШІ. Як зазначає Дж. Крайгер, операціоналізація етики ШІ є «багатовимірною справою, яка потребує міцної інфраструктури для етичного оцінювання, що вимагає дій з кількох напрямків одночасно, таких як управління, політика та навчання» [135, с. 356].

Втім, аналіз ефективності Рекомендації виявляє обмеження, притаманні джерелам «м'якого права» у сфері регулювання ШІ. Добровільний характер Рекомендації та відсутність обов'язкових механізмів виконання ставлять під сумнів дієвість її положень при застосуванні розробниками систем ШІ. Як аргументовано

зазначає Д. Шифф, «коли стандарт є необов'язковим, його дотримання може легко перетворитися на процедуру заради «галочки» або в поверхневій аудиту», за яких компанія може «стверджувати про відповідність певним етичним рамкам, водночас мало роблячи для пом'якшення відчутної шкоди» [132, с. 14]. Показово, що навіть механізм моніторингу імплементації Рекомендації є добровільним і не передбачає санкцій за недотримання, що додатково підкреслює структурні межі виключно етичних інструментів регулювання.

Міжнародний союз електров'язку (МСЕ) посідає особливу позицію серед спеціалізованих установ ООН як технічна організація, відповідальна за розробку міжнародних стандартів інформаційно-комунікаційних технологій. Юридична специфіка актів МСЕ у сфері ШІ полягає у їхній двоїтій природі: з одного боку, резолюції Повноважної конференції та Всесвітньої асамблеї зі стандартизації електров'язку є актами вищих органів Союзу, що обов'язкові для внутрішніх структур МСЕ та створюють програмні зобов'язання щодо напрямів діяльності Організації, але не породжують прямих юридичних обов'язків для держав-членів; з іншого боку, технічні стандарти (рекомендації МСЕ-Т), хоча формально є добровільними, стають де-факто обов'язковими внаслідок вимог технічної сумісності, оскільки відхилення від них унеможлиблює інтероперабельність систем. Цей механізм «м'якого примусу» через технічну необхідність підвищує регуляторну ефективність актів МСЕ порівняно з класичними рекомендаціями інших спеціалізованих установ.

Основною платформою МСЕ у сфері ШІ є «Глобальний Саміт ШІ заради добра» [29], ініційований у 2017 році як перший глобальний форум для обговорення практичного застосування технологій ШІ для досягнення ЦСР. Резолюція МСЕ 214, прийнята в Бухаресті у 2022 році, констатує, що «МСЕ співпрацює з понад 40 іншими установами ООН для організації платформи «ШІ заради добра» та запустила Репозиторій ШІ МСЕ, який має на меті визначити практичне застосування ШІ для просування Цілей сталого розвитку» [112, п. (b)], та інструктує Генерального секретаря «зосередити зусилля МСЕ, пов'язані зі штучним інтелектом, на сприянні

надійним екосистемам телекомунікацій/ІКТ з метою підтримки технологій штучного інтелекту» [112, п. 2]. Резолюція 101 Всесвітньої асамблеї зі стандартизації електрозв'язку, прийнята у Нью-Делі у 2024 році, визначила оновлений мандат Сектору стандартизації електрозв'язку МСЕ (МСЕ-Т) щодо ШІ на наступний чотирирічний період, підкресливши «необхідність глобальної співпраці та діалогу між державами-членами, членами секторів та іншими зацікавленими сторонами у вирішенні можливостей та потенційних викликів ШІ щодо його ролі в електрозв'язку, включаючи аспекти надійності» [109].

Практичну діяльність з технічної стандартизації МСЕ здійснює через систему фокус-груп, що працюють над стандартами для конкретних сфер застосування ШІ: від охорони здоров'я (FG-AI4H, спільно з ВООЗ) та управління природними катастрофами (FG-AI4NDM, спільно з ВМО та ЮНЕП) до автономних транспортних засобів (FG-AI4AD, спільно з ЄЕК ООН) та енергоефективності (FG-AI4EE). Як зазначається у технічному звіті фокус-групи з управління природними катастрофами, «ландшафт стандартизації є надзвичайно динамічним з багатьма відповідними зацікавленими сторонами, що працюють над різноманітними темами» [110, с. 9], що робить координацію між МСЕ, ІСО, МЕК та іншими органами критично важливою для уникнення дублювання зусиль та виявлення прогалів у стандартизації. МСЕ також координує створення глобальної Баз даних стандартів ШІ (англ. – AI Standards Exchange Database), що надає доступ до понад 700 стандартів ШІ, класифікуючи їх за галузевими вертикалями та сферами застосування.

Аналіз діяльності МСЕ виявляє проблему фрагментації глобального ландшафту стандартів у сфері ШІ. Множинність стандартів, розроблених як МСЕ, так й іншими міжнародними організаціями (ISO, IEEE, IEC), з використанням відмінної термінології та методологій, створює значний тягар відповідності для розробників ШІ. Д. Шифф констатує: «в основі цієї плутанини лежить зростаючий тягар, що покладається на команди корпоративного комплаєнсу, розробників та аудиторів, оскільки вони намагаються орієнтуватися, визначати пріоритети та

узгоджувати поширені стандарти» [132, с. 10], а самі міжнародні організації зі стандартизації та національні юрисдикції «розробили перехресні, дублюючі або частково несумісні стандарти з незначною координацією, що призводить до організаційної плутанини та паралічу ухвалення рішень щодо того, які стандарти приймати або надавати пріоритет» [132, с. 19]. Відповіддю на цей виклик може стати конвергенція підходів: за оцінкою того ж автора, таке узгодження «не вимагає, щоб кожна країна або організація прийняла єдиний, монолітний стандарт; натомість воно передбачає узгодження фундаментальних принципів та створення взаємозв'язків між структурами для зниження витрат на дотримання, зменшення суперечливих вимог та сприяння транскордонному співробітництву» [132, с. 21]. Перспективним напрямком є формалізація співпраці між органами зі стандартизації на кшталт ISO, IEEE, IEC, яка могла б забезпечити використання уніфікованої термінологічної бази та методології.

Додатковим системним викликом є проблема «темпорального розриву», виявлена та обґрунтована у підрозділі 1.3 цього дослідження. Д. Шифф влучно описує її практичний прояв на рівні спеціалізованих установ: «стандарти розробки потребують років для досягнення консенсусу, ризикуючи майже миттєвим застаріванням, у той час як корпоративні дослідницькі лабораторії ітерують нові архітектури та моделі впровадження протягом лише кількох місяців, створюючи постійну напругу між процедурною легітимністю та практичною реакцією» [132, с. 17]. Як буде показано далі, саме ця проблема спонукає спеціалізовані установи ООН та міжрегіональні організації до вироблення нових моделей нормотворчої діяльності.

Всесвітня організація охорони здоров'я (ВООЗ) є єдиною спеціалізованою установою ООН, що розробила комплексну систему галузевих актів для конкретної сфери застосування технологій ШІ – охорони здоров'я. З позиції правової природи, акти ВООЗ у цій сфері являють собою керівництва, що за юридичною силою поступаються як рекомендаціям Генеральної конференції ЮНЕСКО, так і резолюціям вищих органів МСЄ: вони не приймаються пленарними органами

Організації шляхом голосування держав-членів, а розробляються секретаріатом за участі експертних груп та випускаються від імені Організації як настанови для національних регуляторів. Така правова форма надає їм авторитет провідної міжнародної медичної установи, але не створює навіть квазіюридичних зобов'язань для держав. Як зазначає керівник відділу ШІ у ВООЗ С. Пуджарі, «для того, щоб ШІ мав благотворний вплив на глобальне здоров'я, особливо в країнах з низьким та середнім рівнем доходу, етичні міркування, регулювання, стандарти та механізми управління повинні бути поставлені в центр проектування, розробки та впровадження систем на основі штучного інтелекту» [131, с. 364].

Нормативну базу ВООЗ у сфері ШІ формують декілька взаємопов'язаних документів. «Глобальна стратегія цифрового здоров'я 2020–2025» [116] визнає, що цифрові технології, включаючи ШІ, мають «доведений потенціал для покращення результатів охорони здоров'я», та встановлює, що «цифрове здоров'я повинно бути невід'ємною частиною пріоритетів охорони здоров'я і приносити користь людям у спосіб, що є етичним, захищеним, безпечним, надійним, справедливим та стійким» [116, с. 9]. Керівництво «Етика та управління штучним інтелектом для охорони здоров'я» [113] закріплює шість етичних принципів для медичних систем ШІ: захист автономії людини, за яким «люди повинні зберігати контроль над системами охорони здоров'я та медичними рішеннями» [113, с. 2]; сприяння добробуту та безпеці, яке вимагає, щоб «технології ШІ не завдавали шкоди людям» та задовольняли регуляторні вимоги щодо безпеки, точності та ефективності [113, с. 2]; забезпечення прозорості та пояснюваності [113, с. 3]; відповідальність та підзвітність, за якою «якщо щось йде не так з технологією ШІ, має бути підзвітність» та доступні механізми відшкодування для постраждалих від алгоритмічних рішень [113, с. 3]; забезпечення інклюзивності та справедливого доступу незалежно від соціально-демографічних характеристик, захищених кодексами прав людини [113, с. 3–4]; та сприяння відповідальному й стійкому ШІ [113, с. 4].

Показовою є здатність ВООЗ оперативно реагувати на технологічні зміни: у січні 2024 року, менш ніж через рік після масового поширення систем генеративного

ШІ, організація випустила спеціалізоване керівництво [114] з понад 40 рекомендаціями щодо цих нових технологій, наполягаючи на тому, що «поширення ШІ для охорони здоров'я повинно відбуватися під наглядом урядів та їхніх відповідних регуляторних органів» [131, с. 364]. Окреме керівництво «Регуляторні вимоги до штучного інтелекту для охорони здоров'я» [115] встановлює процедури клінічної валідації, постмаркетингового нагляду, звітування про несправності, ревалідації та кібербезпеки медичних систем ШІ. Водночас ВООЗ визнає брак інфраструктури для оцінювання: «наразі існує брак глобально визнаних рамок для бенчмаркінгу, щоб оцінювати докази, згенеровані при використанні цих пристроїв» [117]. Особливої уваги заслуговують правові виклики, пов'язані з інтеграцією ШІ у медичну практику: за оцінкою С. Пуджарі, «інтеграція ШІ в традиційну медицину створює регуляторні складнощі та питання правової підзвітності», оскільки «визначення відповідальності у випадках, пов'язаних з помилками ШІ або несприятливими наслідками, може бути складним, що вимагає чітких правових рамок» [130, с. 739].

У 2018 році ВООЗ та МСЕ створили Фокусну групу з штучного інтелекту для охорони здоров'я, яка «об'єднала понад 100 зацікавлених сторін для розробки системи орієнтирів, яка спрямовує проектування, розробку, регулювання та впровадження цих інструментів» [131, с. 364]. ВООЗ закликає до інклюзивної співпраці через багатостороннє партнерство для забезпечення того, щоб інтеграція ШІ в медицину була етичною та чутливою до місцевих і глобальних пріоритетів охорони здоров'я [130, с. 740], та наголошує, що «компанії, що розробляють штучний інтелект, пов'язаний зі здоров'ям, слід заохочувати діяти як відповідальні розпорядники громадського здоров'я, ставлячи добробут та безпеку людей вище комерційних інтересів» [131, с. 364].

Можемо зазначити, що ВООЗ демонструє найбільш розвинений серед спеціалізованих установ ООН секторальний підхід до регулювання ШІ, що поєднує етичні принципи, процедурні вимоги та оперативну адаптацію до технологічних змін. Водночас відсутність обов'язкової сили її керівництв та визнання самою

Організацією брак глобальних рамок для бенчмаркінгу медичних систем ШІ засвідчують, що навіть найбільш розроблена секторальна модель залишається в межах «м'якого права» і не може самостійно забезпечити ефективний захист пацієнтів від ризиків, пов'язаних із застосуванням ШІ в охороні здоров'я.

Всесвітня організація інтелектуальної власності (ВОІВ) посідає окрему нішу в системі регулювання ШІ, оскільки її діяльність зосереджується не на встановленні стандартів поведінки, а на картографуванні правових проблем, породжених технологіями ШІ для усталеної системи інтелектуальної власності. З позиції юридичної природи, акти ВОІВ у сфері ШІ принципово відрізняються від актів інших спеціалізованих установ: організація свідомо утримується від вироблення нормативних рекомендацій чи керівництв, натомість продукуючи аналітичні звіти, проєкти документів з питань політики та матеріали дискусійних платформ, що мають характер попереднього дослідження (англ. – *fact-finding*), яке передуює нормотворчості. Вони не кваліфікуються навіть як «м'яке право», оскільки не містять нормативних приписів, адресованих державам. Причина такої стриманості полягає у фундаментальній невизначеності правових питань на перетині ШІ та інтелектуальної власності, де передчасна фіксація позицій могла б завдати шкоди інтересам як правовласників, так і розробників ШІ.

Звіт «Технологічні тренди штучного інтелекту» [118], опублікований у 2019 році, став першим масштабним емпіричним аналізом патентної активності у сфері ШІ. За даними звіту, з моменту появи ШІ було подано заявки на майже 340 тисяч винаходів, причому понад половина з них – після 2013 року, при цьому домінуючою технологією є машинне навчання, яке «включається більш ніж в одну третину всіх ідентифікованих винаходів» [118], а патентна активність концентрується у США та Китаї, що разом з Японією становлять «78% від загальної кількості патентних заявок» [118]. Генеральний директор ВОІВ Ф. Гаррі констатував, що «штучний інтелект є новим цифровим кордоном, який матиме глибокий вплив на світ, трансформуючи спосіб, яким ми живемо та працюємо» [118, с. 3], та наголосив на необхідності оперативної політичної реакції на стрімкі технологічні зміни.

Діалог ВОІВ щодо інтелектуальної власності та штучного інтелекту, розпочатий у 2019 році, створив платформу для обговорення правових питань на перетині технологій ШІ та інтелектуальної власності. Відкриваючи першу сесію, Генеральний директор ВОІВ зазначив, що ШІ є «відносно новою» технологією загального призначення, яка «має та матиме надзвичайно глибокий вплив» на економіку та суспільство, водночас попередивши: «існують явні лідери і є ті, хто відстає. Необхідно усвідомлювати наявність великого технологічного розриву у світі, і ШІ ризикує посилити цей розрив» [121]. У грудні 2019 року ВОІВ оприлюднила «Проект документа з питань політики інтелектуальної власності та штучного інтелекту» [119], на який було отримано понад 250 коментарів від держав-членів та приватних організацій. Серед ключових правових питань, ідентифікованих ВОІВ, виокремлюються дві проблемні групи: питання винахідництва та авторства (чи може система ШІ бути визнана винахідником, як розподіляти права між розробниками, користувачами та власниками навчальних даних) та питання використання захищених творів для навчання систем ШІ (чи є таке використання порушенням авторського права, як застосовуються винятки щодо добросовісного використання, яку компенсацію мають отримувати правовласники). Керівництво «Генеративний штучний інтелект: навігація інтелектуальною власністю» констатує, що «існує значна правова невизначеність щодо того, чи представляють інструменти ШІ, їхнє навчання, використання та результати порушення прав інтелектуальної власності» [120], і що «незрозуміло, чи може новий контент, згенерований інструментами ШІ... бути захищений правами інтелектуальної власності, і якщо так, то хто володіє цими правами» [120]. Додатково керівництво попереджає про ризик ненавмисного розкриття комерційних таємниць при використанні інструментів генеративного ШІ та фіксує наявність «кількох поточних судових спорів, що стверджують, що збір та використання цих творів для навчання ШІ... є порушеннями інтелектуальної власності» [120], причому «відповідь може відрізнятись залежно від юрисдикції» [120]. ВОІВ також створила Глобальну базу даних національних політик у сфері інтелектуальної власності та штучного інтелекту [122], що дозволяє

відстежувати фрагментацію правових підходів різних юрисдикцій. Отже, ВОІВ займає унікальну позицію серед спеціалізованих установ ООН: замість вироблення нормативних приписів вона виконує функцію «правового картографа», систематизуючи проблемне поле на перетині ІІІ та інтелектуальної власності.

Міжнародна організація праці (МОП) вносить у систему регулювання ІІІ внесок, якісно відмінний від усіх інших спеціалізованих установ: замість формулювання стандартів поведінки чи етичних принципів, організація забезпечує емпіричну доказову базу для обґрунтування необхідності регуляторного втручання. Акти МОП у сфері ІІІ являють собою робочі документи (англ. – *working papers*) та аналітичні звіти, підготовлені дослідницькими підрозділами Організації, і не ухвалюються Міжнародною конференцією праці чи Адміністративною радою МОП. Їхня цінність полягає не в нормативній силі, а у створенні доказової бази, що інформує та легітимізує подальші нормотворчі ініціативи як самої МОП, так і інших спеціалізованих установ ООН.

Робочий документ «Генеративний штучний інтелект та робочі місця: глобальний аналіз» [123], опублікований у 2023 році, виявив, що з усіх професійних груп найвища експозиція до технології спостерігається в канцелярській роботі, де «24 відсотки канцелярських завдань вважаються високо експонованими, та додаткові 58 відсотків із середнім рівнем експозиції» [123, с. 1], тоді як для інших професійних груп частка високо експонованих завдань становить лише 1–4 відсотки. Дослідження зафіксувало також різку географічну асиметрію: «у країнах з низьким рівнем доходу лише 0,4 відсотка від загальної зайнятості потенційно експоновані до ефектів автоматизації, тоді як у країнах з високим рівнем доходу частка зростає до 5,5 відсотка» [123, с. 1]. Значущим для міжнародно-правового регулювання є виявлена МОП гендерна диспропорція: «ефекти є високо гендеризованими, з понад подвоєнням частки жінок, потенційно уражених автоматизацією» [123, с. 1], при цьому 9,6% жіночої зайнятості у світі знаходиться в роботах, потенційно автоматизованих за допомогою ІІІ, порівняно з 3,5% чоловічої зайнятості [124]. Організація застерігає, що «концентровані втрати робочих місць у професіях, де

домінують жінки, можуть загрожувати досягненням, зробленим у минулі десятиліття у збільшенні участі жінок на ринку праці» [123]. Регіональне дослідження впливу генеративного ШІ на ринки праці Латинської Америки підтвердило критичну роль цифрового розриву: «майже половина робочих місць, які потенційно могли б отримати вигоду від доповнення, обмежені через відсутність використання цифрових технологій» [126, с. 1] – йдеться приблизно про 17 мільйонів робочих місць у 16 країнах регіону [126, с. 29].

У 2025 році МОП представила два нових дослідження. «Удосконалений глобальний індекс професійної експозиції до генеративного ШІ» [124], підготовлений у партнерстві з NASK, оновив методологію з огляду на розвиток великих мовних моделей з можливостями розпізнавання зображень та виробництва аудіо та відео і базується на аналізі трудових функцій понад 1000 професій. Дослідження «Робота трансформована: обіцянки та небезпеки штучного інтелекту» [125] виявило, що «замість того, щоб повністю замінювати людську працю, ШІ може доповнювати та покращувати її, звужуючи розрив між недосвідченими та експертними працівниками» [125]. МОП також створила Обсерваторію штучного інтелекту та ринку праці [127] для постійного моніторингу впливу технологій ШІ на зайнятість.

Можемо зазначити, що МОП фактично виконує функцію емпіричного фундатора: її дослідження створюють доказову базу, без якої нормотворчі ініціативи інших установ системи ООН не мали б достатнього емпіричного обґрунтування. Водночас сама МОП досі не перейшла від дослідницької до нормотворчої діяльності у сфері ШІ: жоден із проаналізованих актів не містить стандартів, конвенцій чи рекомендацій, ухвалених Міжнародною конференцією праці, що залишає соціально-трудовий вимір регулювання ШІ без спеціалізованого міжнародно-правового акту.

Програма розвитку ООН (ПРООН), хоча й не є спеціалізованою установою у формально-правовому розумінні, відіграє роль практичного імплементації ініціатив, розроблених спеціалізованими установами ООН. Документи ПРООН у сфері ШІ – стратегії, проєктні документи та методичні посібники – мають

внутрішньоорганізаційний характер та не адресуються безпосередньо державам як суб'єктам міжнародного права. Цифрова стратегія ПРООН на період 2022–2025 років формулює візію «допомогти створити світ, у якому цифрові технології є силою, що надає можливості людям та планеті» [129, с. 3], визначаючи ШІ частиною ширшого спектру цифрових технологій [129, с. 5]. Стратегія закріплює людиноцентричний підхід: «права людини будуть у центрі підходу ПРООН до цифрових технологій, оскільки цифрові технології все більше опосередковують все – від охорони здоров'я, зайнятості, освіти до участі у соціальному, культурному та політичному житті» [129, с. 10]. Унікальним інструментом ПРООН є мережа *Accelerator Labs*, що охоплює 115 країн, підтримує місцеві інновації та створює адаптовані до локального контексту рішення [129, с. 13, 36]. Організація зобов'язується «допомагати партнерам ПРООН ефективно керувати цифровими технологіями, включаючи потенційні ризики, пов'язані з кібербезпекою, приватністю даних, дезінформацією та іншими сферами» [129, с. 24].

Окремим елементом регулювання ШІ в рамках системи ООН є Біла книга системи ООН з управління штучним інтелектом [128], підготовлена Міжвідомчою робочою групою зі штучного інтелекту у 2024 році. Ця робоча група, «співочолювана МСЕ та ЮНЕСКО та об'єднуючи понад 40 установ ООН» [128, п. 20], здійснила масштабне міжвідомче дослідження, що включало опитування всієї системи ООН, інтерв'ю з вибраними установами та аналіз глобальних тенденцій і існуючої літератури. Книга надає огляд «поточних інституційних моделей та існуючих міжнародних нормативних рамок системи ООН, які можуть бути застосовані або використані для міжнародного управління штучним інтелектом» [128, п. 15], та підкреслює центральне місце системи ООН, що «грає унікальну роль у встановленні норм та досягненні консенсусу, розбудові спроможності та міжнародному співробітництві в питаннях, що охоплюють права людини, клімат, роззброєння, здоров'я та технології, і готова бути використаною для подолання нових викликів або нових сфер управління» [128, п. 13]. Водночас Біла книга визнає структурні обмеження, наголошуючи на необхідності «враховувати швидкий темп

технологічного розвитку, порівняно з відносно повільним процесом розробки нових інструментів міжнародного права в інституційних структурах, потребу в регіональних або галузевих підходах, а також рівень гнучкості механізмів та процесів, що існують на інституційному рівні» [128, п. 9(е)].

Діяльність спеціалізованих установ ООН тісно пов'язана з резолюціями Генеральної Асамблеї через систему прямих мандатів та концептуальної синергії. Резолюція ГА ООН A/RES/78/265 у преамбулі посилається на прийняття Генеральною конференцією ЮНЕСКО Рекомендації про етичні аспекти штучного інтелекту та плану її здійснення, а в п. 10 «закликає спеціалізовані установи, фонди, програми, інші структури, органи та установи і відповідні організації системи ООН продовжувати, в рамках їхніх відповідних мандатів та ресурсів, оцінювати та посилювати вжиті ними заходи реагування з метою використання можливостей та вирішення проблем, створюваних системами штучного інтелекту, на спільній, скоординованій та всеохоплюючій основі» [19, п. 10]. Резолюція A/RES/79/1 «Пакт заради майбутнього» [21], що включає Глобальний цифровий договір, у п. 52 безпосередньо посилається на Рекомендацію ЮНЕСКО про етичні аспекти ШІ як відповідний рамковий акт, а в пп. 55 та 61 встановлює мандат для координованої діяльності через існуючі механізми ООН, зокрема зобов'язуючись «використовувати існуючі механізми ООН та багатосторонні механізми для сприяння нарощуванню потенціалу в галузі штучного інтелекту» [21, п. 61]. Резолюція A/RES/78/311 у п. 9 надає конкретний мандат спеціалізованим установам системи ООН «здійяти можливості та вирішувати проблеми, пов'язані зі зміцненням міжнародного співробітництва у справі нарощування потенціалу в галузі штучного інтелекту», а в п. 2 рекомендує «збільшити фінансування та технічну допомогу країнам, що розвиваються, в галузі нарощування потенціалу» [86, пп. 2, 9]. Міжвідомча робоча група зі штучного інтелекту, спільні фокус-групи та система перехресних посилянь у нормативних актах різних установ створюють багаторівневу мережу координації, а платформа «Глобальний Саміт ШІ заради добра» щорічно збирає представників понад 40 установ системи ООН.

Незважаючи на значні досягнення, аналіз виявляє низку критичних прогалин у регуляторному покритті спеціалізованих установ ООН. Найбільш суттєвою є відсутність повного мандату будь-якої установи на регулювання воєнного використання технологій ШІ. Хоча ГА ООН ухвалила резолюції A/RES/80/57 [94] та A/RES/80/58 [95] щодо автономних систем озброєнь із ШІ, жодна спеціалізована установа не має відповідних імплементаційних механізмів. На відсутність цілісного договірного регулювання у сфері воєнного застосування ШІ звертає увагу й вітчизняна доктрина. Як зазначає М. О. Медведєва, у цій сфері «не існує єдиного універсального договору, який би регулював ведення війни в епоху штучного інтелекту» [289, с. 187], а наявні акти – резолюції Генеральної Асамблеї ООН та керівні принципи Групи урядових експертів – мають рекомендаційний характер. Це узгоджується з виявленою нами закономірністю щодо воєнного використання ШІ, яке, попри ухвалення серії резолюцій, не підкріплене обов'язковими механізмами контролю. Самі ризики застосування ШІ у воєнній сфері у попередніх авторських дослідженнях запропоновано розмежовувати на дві групи – технічні та етично-правові: до перших належать помилки системи, хибна ідентифікація цілей та вразливість до кібератак, до других – «використання автономних бойових систем, які приймають рішення без участі людини, що створює ризики застосування сили без належного розуміння контексту або моральних наслідків» [108, с. 300], а також неможливість встановлення суб'єкта відповідальності за дії ШІ. Найгостріше серед етично-правових ризиків постає недотримання норм міжнародного гуманітарного права: технічні обмеження автономних систем означають, що вони наразі не здатні повною мірою дотримуватися принципу розрізнення, закріпленого в статті 48 Додаткового протоколу I до Женевських конвенцій, через «семантичний розрив» між алгоритмічним розпізнаванням об'єктів та розумінням їх значення в контексті збройного конфлікту [62, с. 155]. У нашому дослідженні з зазначеної проблематики [62] обґрунтовано трикомпонентну модель забезпечення дотримання принципу розрізнення при використанні автономних систем озброєння із ШІ, що включає створення спеціалізованої міжнародної агенції за моделлю МАГАТЕ – Міжнародної

агенції з контролю автономних систем озброєння (МАКАСО) [62, с. 155], трирівневу процедуру сертифікації автономних систем із ШІ з обов'язковим аудитом алгоритмів незалежними експертами та постійним моніторингом через систему «чорних скриньок» [62, с. 156], а також проєкт «Конвенції про тимчасову заборону повністю автономних систем із ШІ», яка має встановлювати заборону на системи категорії Human-out-of-the-Loop, зберігаючи допустимість систем Human-in-the-Loop та Human-on-the-Loop зі значущим контролем людини [62, с. 156].

Іншими значними прогалинами є недостатнє регулювання екологічного впливу технологій ШІ (жодна установа не розробила стандартів щодо енергоспоживання дата-центрів чи карбонового сліду навчання моделей), фрагментарне покриття питань управління даними (ЮНЕСКО розглядає етичні аспекти, ВООЗ – медичні дані, ВОІВ – дані як об'єкт інтелектуальної власності, але комплексна міжнародна рамка відсутня), недостатня увага до кіберзлочинності з використанням ШІ (дипфейки, автоматизовані кібератаки, маніпуляції виборчими процесами), а також ігнорування проблематики штучного загального інтелекту (AGI) та довгострокових екзистенційних ризиків. Недостатньо розвиненими залишаються й механізми верифікації імплементації: моніторинг ЮНЕСКО є добровільним, постмаркетинговий нагляд ВООЗ залежить від спроможності національних регуляторів, МОП не має механізмів перевірки адекватності заходів держав-членів, а сертифікація МСЕ застосовується переважно до обладнання, а не до алгоритмів.

На підставі індуктивного узагальнення проаналізованих актів та ініціатив спеціалізованих установ ООН можна ідентифікувати п'ять закономірностей, що характеризують формування міжнародно-правового регулювання ШІ на цьому рівні. Першою є формування глобального ціннісного консенсусу: незважаючи на різні мандати, всі установи демонструють узгодженість позицій щодо базових цінностей – прав людини, прозорості, справедливості та інклюзивності, причому Рекомендація ЮНЕСКО [13] функціонує як етичний фундамент, на який спираються ініціативи ВООЗ, МСЕ та МОП у їхніх відповідних секторальних мандатах. Другою

закономірністю є безумовне домінування інструментів «м'якого права»: жоден із проаналізованих актів не створює юридично обов'язкових зобов'язань у класичному розумінні міжнародного права, утворюючи спектр з різним ступенем нормативного впливу – від актів з відносно високим рівнем зобов'язання (Рекомендація ЮНЕСКО, прийнята консенсусом 193 держав) до суто аналітичних документів (звіти ВОІВ та МОП). Л. Шмітт пояснює це тим, що «уряди вирішили передати певну владу міжнародним організаціям лише доти, доки вони мають справу з досить абстрактними принципами чи м'яким управлінням, але відступили б або затягували, як тільки робота просунеться до більш регуляторного, твердого управління» [142, с. 310]. Третьою закономірністю є здатність установ до прискореної адаптації до технологічних змін, що контрастує з традиційною повільністю міжнародних інституційних процесів: ВООЗ оновила керівництва для великих мультимодальних моделей менш ніж за рік після їх масового поширення, а МОП і МСЕ оперативно скоригували свої аналітичні та стандартизаційні програми. Четвертою закономірністю є міжінституційна координація через Міжвідомчу робочу групу зі ШІ, спільні фокус-групи МСЕ-ВООЗ та систему взаємних посилянь, яка, проте, залишається недостатньою у горизонтальних сферах – управлінні даними, кібербезпеці, воєнному використанні ШІ, – де кожна установа розглядає лише фрагмент проблеми за відсутності інтегруючого органу. П'ятою закономірністю є стійка географічна та гендерна нерівність у впливі технологій ШІ: емпіричні дані МОП засвідчують суттєвий розрив експозиції робочих місць до автоматизації між країнами з високим і низьким рівнем доходу, а попри зусилля установ системи ООН щодо подолання цифрового розриву та забезпечення інклюзивності, масштаб проблеми суттєво перевищує наявні інституційні ресурси.

Резюмуючи, зазначимо, що діяльність спеціалізованих установ ООН формує багатовимірну систему регулювання ШІ, що поєднує етичні принципи ЮНЕСКО, технічні стандарти МСЕ, галузеві керівництва ВООЗ, аналітичні дослідження ВОІВ та МОП, а також практичний досвід ПРООН. З точки зору правової кваліфікації ця система являє собою приклад того, що в науковій літературі визначається як

«слабкий режимний комплекс» [150, с. 3] – поліцентричний інституційний ландшафт без чіткої ієрархії, в якому акти різного нормативного рівня співіснують та взаємно підсилюють одне одного. Ця система характеризується як значними досягненнями у формуванні глобального консенсусу та адаптації до технологічних змін, так і суттєвими прогалинами в регулюванні воєнного використання ШІ, екологічного впливу, управління даними та довгострокових ризиків. Взаємодія між резолюціями Генеральної Асамблеї ООН та ініціативами спеціалізованих установ створює синергію, при якій загальні принципи, встановлені на рівні ГА ООН, конкретизуються через спеціалізовані акти установ, проте для формування дійсно ефективної системи регулювання необхідне посилення координації, заповнення виявлених прогалин та розробка механізмів забезпечення імплементації.

2.3. Роль ОЕСР та міжрегіональних ініціатив у формуванні міжнародних стандартів застосування штучного інтелекту

Аналіз міжнародно-правового регулювання ШІ, здійснений у попередніх підрозділах на прикладі ГА ООН та спеціалізованих установ системи ООН, засвідчив формування нормативного масиву на універсальному рівні міжнародного співробітництва, для якого характерні максимальна інклюзивність учасників, але водночас і неминучий високий рівень абстрактності правових приписів. Однак поряд з універсальним та регіональним рівнями, які є традиційними для доктрини міжнародного публічного права, у сфері регулювання ШІ особливого значення набуває проміжний рівень, який у цьому дослідженні визначається як міжрегіональний. Саме на цьому рівні зосереджена значна частина нормотворчої активності, пов'язаної із формуванням стандартів відповідального розвитку та використання технологій ШІ.

Під міжнародно-правовим регулюванням ШІ на міжрегіональному рівні у цьому дослідженні розуміється сукупність нормотворчих процесів, що здійснюються міжнародними організаціями, міжурядовими форумами та

ініціативами, які за своїм складом охоплюють держави з різних географічних регіонів, проте не претендують на універсальне членство і не обмежені географічним мандатом конкретного регіону. Віднесення тієї чи іншої інституції до міжрегіонального рівня здійснюється за сукупністю критеріїв: по-перше, міжконтинентальний склад учасників, що виключає замкненість у межах одного географічного регіону; по-друге, відсутність формального географічного обмеження мандату, на відміну від регіональних організацій на кшталт Ради Європи чи Африканського Союзу; по-третє, переважна орієнтація на економічне, технологічне або науково-технічне співробітництво; по-четверте, домінування інструментів «м'якого права» у нормотворчій діяльності. На підставі цих критеріїв до міжрегіонального рівня регулювання ШІ належать насамперед Організація економічного співробітництва та розвитку (ОЕСР), Група семи (G7), Група двадцяти (G20) та Глобальне партнерство зі штучного інтелекту (GPAI).

Доктринальне обґрунтування виокремлення міжрегіонального рівня як самостійного об'єкта аналізу спирається на концепцію поліцентризму, при якій існує багато різних центрів правотворення, суб'єктів та регуляторних систем (міжнародні організації, держави, регіональні союзи). П. Сігон, М. Маас та Л. Кемп у дослідженні архітектури міжнародно-правового регулювання ШІ констатують, що «ШІ має потенціал кардинально змінити світ на краще чи на гірше. Ці високі ставки спричинили нещодавній сплеск міжнародної нормотворчості у сфері ШІ в ОЕСР, Групі семи, Групі двадцяти та численних інституціях ООН» [148, с. 546]. Цей висновок підтверджує, що саме на перетині універсального та регіонального рівнів, у просторі міжрегіональних ініціатив, відбувається динамічна нормотворча діяльність, результати якої згодом транслюються як на універсальний, так і на регіональний та національні рівні. Як зазначає Л. Шмітт, «зароджуваний режим ШІ є поліцентричним і фрагментованим, але гравітує навколо ОЕСР, яка володіє значним епістемічним авторитетом та нормотворчою силою» [142, с. 311]. Ця гравітація навколо ОЕСР, що поширюється на пов'язані з нею формати G7, G20 та

GPAI, формує своєрідний нормативний кластер, аналіз якого є предметом цього підрозділу.

Взаємопов'язаність інституцій міжрегіонального рівня є їхньою визначальною структурною ознакою. ОЕСР виступає центральною ланкою цієї мережі: Принципи ОЕСР зі ШІ 2019 року стали основою для Принципів G20 зі ШІ, прийнятих того ж року на саміті в Осаці [10]; секретаріат GPAI розміщено при ОЕСР; а Хіросімський процес G7 зі ШІ 2023 року реалізовувався за технічної підтримки Організації. Дослідники П. Сігон, М. Маас та Л. Кемп відзначають, що фрагментоване членство у цих ініціативах, де лише частина держав бере участь одночасно в усіх форматах, створює характерну для режимних комплексів картину перехресного членства та тематичного перекриття [148, с. 548]. Подальший аналіз здійснюватиметься від центральної інституції цього кластеру – ОЕСР – до суміжних форматів G7, G20 та GPAI, з акцентом на юридичну природу прийнятих актів, механізми їх імплементації та внесок кожної ініціативи у формування міжнародних стандартів застосування ШІ.

Організація економічного співробітництва та розвитку відіграє особливу роль у формуванні міжнародних стандартів регулювання ШІ, виступаючи одночасно як розробник міжнародних стандартів, координатор регуляторних підходів та платформа для діалогу за участю багатьох зацікавлених сторін. У контексті фрагментованого ландшафту регулювання ШІ, при якому відсутня єдина централізована інституція та діє множина паралельних ініціатив різних акторів, ОЕСР демонструє ознаки консолідуючого центру, навколо якого формується поліцентричний режим міжнародно-правового регулювання ШІ. Л. Шмітт підкреслює, що ОЕСР завдяки «своєму епістемічному авторитетові та нормо- і агендотворчій силі змогла діяти як референтна точка для G7 та G20» [142, с. 311]. Ця характеристика відображає унікальне становище Організації у системі міжнародних відносин, в якій вона, не маючи формальної регуляторної влади чи можливостей примусу, здійснює вплив через технічну експертизу, створення стандартів та сприяння конвергенції національних політик. Такий авторитет ОЕСР у сфері регулювання ШІ ґрунтується на кількох взаємопов'язаних факторах. По-

перше, це багаторічний досвід Організації у розробці міжнародних стандартів у суміжних сферах цифрової економіки, зокрема Керівних принципів щодо захисту приватності та транскордонних потоків персональних даних 1980 року, які стали основою національних законодавств про захист даних у багатьох країнах світу. По-друге, спроможність ОЕСР залучати провідних експертів з різних дисциплін та секторів – урядовців, представників індустрії, громадянського суспільства, профспілок, технічної спільноти та наукових установ – створює унікальну платформу для вироблення консенсусних позицій за участю багатьох зацікавлених сторін. По-третє, методологічна послідовність та аналітична глибина досліджень ОЕСР надають її рекомендаціям високу наукову значущість. Нарешті, інституційні зв'язки з іншими ключовими акторами міжнародних відносин, серед іншого з G7, G20, Європейською Комісією, установами системи ООН та Радою Європи, дозволяють ОЕСР виступати як проміжна ланка між різними рівнями та форматами міжнародного співробітництва.

Втім, слід визнати, що нормотворча роль ОЕСР не є безальтернативною чи всеохопною. Організація об'єднує переважно розвинені країни з високим рівнем доходу, що породжує питання щодо універсальності та репрезентативності напрацьованих нею положень та стандартів. Британські дослідники Г. Робертс, Е. Гайн, М. Таддео та Л. Флоріді застерігають, що «ключовий ризик ОЕСР полягає у певному дефіциті легітимності через те, що вона є переважно європоцентричною організацією, що може призводити до ігнорування її напрацьовань деякими державами» [143, с. 1285]. Це спостереження вказує на фундаментальну напругу між технічною експертизою та політичною легітимністю на міжрегіональному рівні регулювання ШІ.

Рекомендація Ради ОЕСР зі штучного інтелекту [9] є результатом поступової еволюції усвідомлення міжнародною спільнотою необхідності узгодженого підходу до регулювання ШІ. Робота ОЕСР у цій сфері розпочалася у 2016 році з Форуму технологічного прогнозування зі штучного інтелекту, була продовжена у 2017 році міжнародною конференцією «Штучний інтелект: розумні машини, розумна

політика» та супроводжувалася масштабною аналітичною роботою з картування технічних, економічних та політико-правових вимірів технології. Як зазначається у самій Рекомендації, «ця робота продемонструвала необхідність формування стабільного політичного середовища на міжнародному рівні для сприяння довірі та впровадженню ШІ у суспільстві» [9, с. 2]. У травні 2018 року Комітет з політики цифрової економіки ОЕСР ухвалив рішення про створення Неформальної групи експертів зі ШІ у складі понад п'ятдесяти фахівців з різних дисциплін та секторів, на основі вихідного документа якої було розроблено проєкт Рекомендації. Рада ОЕСР прийняла Рекомендацію на засіданні на рівні міністрів 22–23 травня 2019 року, а вже у червні 2019 року на саміті в Осаці лідери G20 привітали Принципи G20 зі ШІ [10], запозичені з Рекомендації ОЕСР, що свідчило про швидке поширення закріплених у ній стандартів поза межами безпосередніх членів Організації.

Подальший розвиток Рекомендації засвідчив спроможність ОЕСР адаптуватися до швидких технологічних змін. Рекомендація пройшла два цикли перегляду. Перший відбувся 8 листопада 2023 року, коли Рада ОЕСР оновила визначення «системи ШІ» для відображення специфіки генеративного ШІ, що набув масового поширення того року. Другий, більш масштабний перегляд Рада здійснила на засіданні міністрів 3 травня 2024 року на основі Звіту про імплементацію, поширення та тривалу актуальність Рекомендації. Оновлена Рекомендація тепер «безпосередньо звертається до викликів, пов'язаних зі ШІ, що стосуються приватності, прав інтелектуальної власності, безпеки та цілісності інформації» [9]. Рекомендацію підтримали сорок сім прихильників, включаючи Європейський Союз, який приєднався 2 травня 2024 року. Така широка адгезія засвідчує міжрегіональне визнання Рекомендації як нормативного орієнтиру у міжнародно-правовому регулюванні ШІ.

З точки зору юридичної природи, Рекомендація Ради ОЕСР є актом «м'якого права», що не створює юридично обов'язкових зобов'язань для держав-членів. Відповідно до статті 5(b) Конвенції про ОЕСР 1960 року, рекомендації Ради подаються на розгляд членів для імплементації, але формально не мають

обов'язкової сили на відміну від рішень Ради. Це визначає специфіку механізму впливу Рекомендації: вона діє не через юридичний примус, а через епістемічний авторитет, соціальний тиск (англ. – *peer pressure*) та поступову інтернаціоналізацію закріплених у ній стандартів через національне законодавство і практику.

Структура та зміст Рекомендації відображають збалансований підхід між встановленням загальних принципів та наданням конкретних рекомендацій для національної політики. Доповнюючи існуючі стандарти ОЕСР, які вже є релевантними для ШІ – такі як стандарти щодо приватності та захисту даних, управління ризиками цифрової безпеки та відповідальної ділової поведінки – Рекомендація зосереджується на політичних питаннях, специфічних для ШІ, і прагне встановити стандарт, який є імplementованим та достатньо гнучким, щоб витримати випробування часом у галузі, що швидко розвивається. Рекомендація містить п'ять високорівневих принципів, заснованих на цінностях, та п'ять рекомендацій для національної політики та міжнародного співробітництва. Вона також пропонує спільне розуміння ключових термінів, таких як «система ШІ», «життєвий цикл системи ШІ» та «актори ШІ» для цілей Рекомендації. Така термінологічна уніфікація є важливим здобутком, оскільки створює спільну мову для міжнародного діалогу та сприяє взаємному розумінню між різними юрисдикціями.

Першу змістовну секцію Рекомендації становлять «Принципи відповідального керівництва надійним ШІ», що встановлюють п'ять взаємодоповнюючих принципів: інклюзивне зростання, сталий розвиток та добробут; повага до верховенства права, прав людини та демократичних цінностей, включно зі справедливістю та приватністю; прозорість та пояснюваність; надійність, безпека та захищеність; підзвітність [9]. Принцип сталого розвитку охоплює і екологічну складову, розгорнуту у методологічному документі ОЕСР 2022 року «Вимірювання екологічних впливів обчислень та застосунків штучного інтелекту: екологічний слід ШІ» [139], що зафіксував значне навантаження великомасштабних систем ШІ на енергетичні та водні ресурси і викиди парникових газів. Принцип людиноцентричних цінностей закріплює інкорпорацію цінностей прав людини,

демократії та верховенства права протягом усього життєвого циклу системи ШІ. Американський дослідник С. Фельдштейн зауважує конвергенцію Рекомендації з Керівними принципами ЄС щодо надійного ШІ: «ключові концепції, такі як справедливість, пояснюваність, прозорість, підзвітність та надійність, є виразними у обох документах... що свідчить про згоду щодо того, як ці цінності мають бути виражені та операціоналізовані» [144, с. 1055]. Принцип прозорості та пояснюваності вимагає від акторів ШІ надавати інформацію, «щоб ті, на кого негативно вплинула система ШІ, могли оскаржити її результат на основі простої та зрозумілої інформації про фактори та логіку, яка слугувала основою для прогнозу, рекомендації чи рішення» [9], а принципи надійності та підзвітності встановлюють вимоги до запровадження запобіжних механізмів і покладають відповідальність за належне функціонування систем ШІ на її акторів [137, с. 7].

Другу змістовну секцію Рекомендації становлять рекомендації для національної політики та міжнародного співробітництва. Прихильники закликаються інвестувати у відповідальні дослідження та розробку ШІ, сприяти інклюзивній цифровій екосистемі, формувати сприятливе інтероперабельне управлінське середовище, готуватися до трансформації ринку праці та сприяти міжнародному співробітництву для надійного ШІ. П'ята рекомендація визнає глобальну природу технологій ШІ та необхідність узгоджених міжнародних дій, закликаючи до обміну найкращими практиками, стандартами та даними, а також підкреслює роль міжнародних організацій, включно з самою ОЕСР, у координації глобальної політики щодо регулювання ШІ.

Імплементация Рекомендації забезпечується через інституційну інфраструктуру та практичні інструменти ОЕСР. Центральну роль відіграє Обсерваторія політики ОЕСР зі штучного інтелекту (англ. – *OECD.AI Policy Observatory*) [8], ініційована у 2020 році як платформа для обміну інформацією про національні політики у сфері ШІ: за даними ОЕСР, «станом на травень 2023 року 51 країна повідомила про наявність національної стратегії зі ШІ у базі даних національних AI-політик OECD.AI» [138, с. 11]. У 2022 році ОЕСР створила Робочу

групу з управління ШІ (англ. – *Working Party on AI Governance*, далі – AIGO) [136] як спеціалізований орган для імплементації Рекомендації та адаптації її положень до нових викликів, разом з Мережею експертів ОЕСР, що об'єднує сотні фахівців з усього світу. Серед розроблених практичних інструментів варто відзначити посібник 2021 року «Інструменти для надійного ШІ: рамка» [137], що систематизував технічні, процедурні та освітні інструменти і має на меті «структурування інформації та полегшення порівняння між різними інструментами» [137, с. 12], «Рамку ОЕСР для класифікації систем ШІ» 2022 року [57], дослідження регуляторних пісочниць 2023 року [140] та аналіз 2024 року «ШІ та приватність: Об'єднуючи Керівні принципи ОЕСР щодо приватності та Принципи ШІ» [141], який констатує: «глобальне значення Принципів ОЕСР зі ШІ неможливо переоцінити. З моменту їх прийняття країни по всьому світу вживали заходів для кодифікації Принципів через національні стратегії ШІ, м'яке та тверде право» [141].

Поряд з ОЕСР, вагому роль у формуванні міжнародних стандартів застосування ШІ на міжрегіональному рівні відіграє Група семи (G7), яка виступає політичним каталізатором нормотворчих процесів. На відміну від ОЕСР, G7 не є міжнародною організацією у формально-юридичному розумінні: вона не має установчого договору, постійного секретаріату чи правосуб'єктності за міжнародним правом. Це неформальний міжурядовий форум, акти якого мають виключно політичний характер і не породжують юридичних зобов'язань для учасників. Проте саме ця інституційна гнучкість дозволяє G7 оперативно реагувати на нові виклики та формувати політичний порядок денний, який згодом трансформується у нормативні документи в рамках формальних міжнародних організацій.

Залучення G7 до проблематики ШІ розпочалося у 2017 році, коли під час головування Канади було ухвалено спільне бачення майбутнього штучного інтелекту – «Шарлевуаську спільну візію», що містить дванадцять зобов'язань з просування людиноцентричного ШІ, який має сприяти економічному зростанню, суспільній довірі та рівності. Згодом, під час головування Японії у 2023 році, роль

G7 у регулюванні ШІ набула якісно нового виміру завдяки ініціюванню Хіросімського процесу зі штучного інтелекту (англ. – *Hiroshima AI Process*). У квітні 2023 року на зустрічі міністрів цифрових технологій G7 у Такасаці було погоджено міністерську декларацію, що наголосила на важливості міжнародних дискусій щодо інтеперабельності різних рамок регулювання ШІ та інвентаризації можливостей і викликів генеративного ШІ. Міністерська дискусія була ескальована до рівня лідерів країн на саміті G7 у Хіросімі у травні 2023 року, де вони доручили міністрам заснувати Хіросімський процес для продовження обговорення генеративного ШІ «в інклюзивний спосіб та у співпраці з ОЕСР та GPAI» [276, п. 26].

Робочі засідання в рамках Хіросімського процесу, що проходили за підтримки ОЕСР з червня по вересень 2023 року, завершилися прийняттям у жовтні 2023 року двох ключових документів: Міжнародних керівних принципів Хіросімського процесу для організацій, що розробляють передові системи ШІ [145], та Міжнародного кодексу поведінки Хіросімського процесу для організацій, що розробляють передові системи ШІ [147]. Керівні принципи містять одинадцять положень, що охоплюють зобов'язання щодо вжиття належних заходів протягом розробки передових систем ШІ, ідентифікації та пом'якшення ризиків, забезпечення прозорості та підзвітності, обміну інформацією про інциденти, інвестування у кібербезпеку, маркування контенту, створеного ШІ, та сприяння дослідженням у сфері безпеки. Кодекс поведінки конкретизує ці принципи та визначає добровільні зобов'язання для розробників передових систем ШІ. У грудні 2023 року лідери G7 погодили Комплексну рамку політики Хіросімського процесу зі ШІ, що об'єднала звіт ОЕСР щодо спільного розуміння генеративного ШІ [149], керівні принципи, кодекс поведінки та проєктне співробітництво у сфері ШІ.

Юридична природа актів Хіросімського процесу заслуговує на окрему характеристику. З формально-правової точки зору, ці акти є політичними зобов'язаннями, що не мають юридично обов'язкової сили. Вони не потребують ратифікації, не передбачають механізмів примусового виконання і не створюють прав чи обов'язків для третіх сторін. Проте їхнє значення полягає у кількох аспектах.

По-перше, вони конкретизують загальні принципи ОЕСР щодо конкретного типу систем ШІ – передових генеративних моделей – заповнюючи нормативну прогалину, яку Рекомендація ОЕСР залишала відкритою. По-друге, політична вага G7 як об'єднання провідних промислово розвинених демократій надає цим документам значну практичну дієвість: великі технологічні компанії, що базуються переважно у країнах G7, сприймають ці зобов'язання як орієнтир для своєї діяльності. По-третє, Хіросімські принципи та Кодекс поведінки спрямовані безпосередньо до організацій-розробників ШІ, а не лише до держав, що є відносно новим підходом на міжрегіональному рівні регулювання та відображає визнання центральної ролі приватного сектору у розвитку ШІ.

Група двадцяти (G20) забезпечує ширший форум для міжнародного діалогу щодо ШІ, об'єднуючи як розвинені країни, так і провідні економіки, що розвиваються, включно з Китаєм, Індією та Бразилією. Це надає актам G20 у сфері ШІ більшу репрезентативність порівняно з актами G7. У червні 2019 року, під японським головуванням, міністри G20 ухвалили Заяву щодо торгівлі та цифрової економіки у Цукубі, в якій схвалили Принципи G20 зі штучного інтелекту, що ґрунтуються на Рекомендації ОЕСР [10]. Прийняття цих принципів лідерами G20 на саміті в Осаці стало, за оцінкою Л. Шмітта, «значним досягненням ОЕСР, яка таким чином розширила свій вплив на деяких ключових гравців поза межами свого членства, зокрема Китай та Росію» [142, с. 306]. Підтримка Принципів G20 зі ШІ Китаєм заслуговує на особливу увагу, адже це засвідчило можливість досягнення мінімального нормативного консенсусу щодо ШІ навіть між геополітичними конкурентами.

Утім, за високорівневою формальною згодою криється суттєвий розрив в інтерпретації узгоджених принципів. Як зауважують Г. Робертс, Е. Гайн, М. Таддео та Л. Флоріді, «високорівневі принципи, такі як погоджені G20, є нечіткими та сумісними з різними ідеологічними позиціями. Візьмімо справедливість у сфері ШІ – принцип, підтриманий усіма членами G20 – у його застосуванні до технології розпізнавання облич. Імплементация цього принципу в ЄС передбачає заборону

таких технологій, тоді як у Китаї технології етнічного розпізнавання є допустимими заради соціальної стабільності» [143, с. 1277]. Ця розбіжність ілюструє фундаментальне обмеження нормативного консенсусу на рівні G20: принципи є достатньо абстрактними для досягнення формальної згоди, але недостатньо конкретними для забезпечення справжньої конвергенції регуляторних підходів. Деякі дослідники, серед них Т. Єлінек, В. Валлах та Д. Керімі, пропонували створення спеціального координаційного комітету G20 з управління штучним інтелектом [151], однак такого прогресу досягнуто не було, що відображає складнощі інституціоналізації співпраці у сфері ШІ в рамках цього неформального форуму.

Глобальне партнерство зі штучного інтелекту (англ. – *Global Partnership on AI*, далі - GPAI) є ще однією інституцією міжрегіонального рівня, що заслуговує на детальний аналіз, – не в останню чергу через радикальну інституційну трансформацію, якої воно зазнало у 2024 році. GPAI було започатковано у 2017 році Канадою та Францією під час їхніх відповідних головувань у G7 під початковою назвою «Міжнародна панель зі штучного інтелекту». Як зазначає Л. Шмітт, «початкова реакція була стриманою, і пропозиція тривалий час нашттовувалась на сильний опір з боку адміністрації Трампа через побоювання, що будь-які кроки до регулювання можуть перешкодити інноваціям у сфері ШІ. Врешті-решт, у травні 2019 року США змінили курс, тепер розглядаючи GPAI як корисний інструмент обмеження впливу Китаю на зароджувану глобальну систему управління ШІ» [142, с. 307]. Офіційний запуск GPAI відбувся у червні 2020 року за участю п'ятнадцяти країн-засновниць, і вже до кінця 2020 року до Партнерства приєдналися Бразилія, Нідерланди, Польща та Іспанія. Первісна інституційна архітектура GPAI поєднувала елементи міжурядової організації та експертної мережі. Партнерство очолювалося Радою на міністерському рівні та Керівним комітетом, а його секретаріат розміщувався при ОЕСР. Практична діяльність зосереджувалася у чотирьох робочих групах за тематичними напрямками: відповідальний ШІ, управління даними, майбутнє праці та інновації і комерціалізація. Ці робочі групи підтримувалися двома

центрами експертизи: Міжнародним центром експертизи з просування ШІ у Монреалі та Французьким національним інститутом досліджень у галузі цифрових наук та технологій (INRIA) у Парижі [28]. Л. Шмітт характеризує інституційну еволюцію GPAI як унікальний випадок: «задумане в рамках існуючої архітектури, воно було запущене як окрема самостійна ініціатива з унікальною членською базою, що виходить за межі G7. Таким чином, воно демонструє характеристики нового самостійного інструменту, водночас врешті-решт опинившись під егідою однієї з існуючих міжнародних організацій (ОЕСР)» [142, с. 307].

Разом з тим ефективність GPAI як самостійного інструменту міжнародно-правового регулювання ШІ залишалася предметом критичної оцінки. Попри амбітну організаційну структуру та широку членську базу, практичні результати діяльності робочих груп обмежувалися переважно дослідницькими звітами, рекомендаціями та пілотними проєктами, що не мають нормативної сили. Г. Робертс, Е. Гайн, М. Таддео та Л. Флоріді зауважують, що «навіть у нових специфічних для ШІ інституціях з вузьким членством, таких як GPAI, було досягнуто мало прогресу у погодженні змістовних ініціатив з регулювання» [143, с. 1277]. Крім того, відсутність у складі GPAI Китаю та Росії обмежувала репрезентативність Партнерства і підривала його претензії на глобальну роль.

Саме ці структурні проблеми – недостатня інституційна стійкість, обмежений бюджет та нечітке розмежування компетенцій із ОЕСР – стали катализатором радикальної інституційної реформи, здійсненої у 2024 році. 3–4 липня 2024 року на Саміті GPAI у Нью-Делі під головуванням Індії було оголошено про створення інтегрованого партнерства GPAI та ОЕСР [277]. Це рішення, розроблене впродовж першої половини 2024 року Спеціальною робочою групою Виконавчої ради GPAI, фактично означало злиття двох паралельних напрямів роботи – програми ОЕСР зі штучного інтелекту та діяльності GPAI – під єдиним брендом GPAI [277]. Інтегроване партнерство об'єднало на засадах рівноправності країни – члени ОЕСР та учасниці GPAI, сформувавши групу з 44 держав з шести континентів, з перспективою розширення за рахунок країн, що розвиваються [277]. Нова

інституційна архітектура передбачає двошарову структуру: управлінський рівень (Рада на міністерському рівні, Пленум як основний робочий орган та Керівна група) та експертний рівень (Експертна спільнота, що об'єднала колишню Мультистейкхолдерну експертну групу GPAI та мережу експертів ОЕСР ONE AI). Секретаріат Партнерства функціонує при ОЕСР та діє за її процедурними правилами [277]. Белградська декларація, прийнята на Саміті GPAI у грудні 2024 року, стала першим актом інтегрованого партнерства, засвідчивши завершення перехідного періоду [278].

Інтеграція GPAI та ОЕСР має принципове значення для проведеного у цій дисертації аналізу у кількох вимірах. По-перше, вона підтверджує тезу про гравітацію поліцентричного режиму міжнародно-правового регулювання ШІ навколо ОЕСР як консолідуючого центру: якщо раніше ця гравітація проявлялася опосередковано – через запозичення Принципів ОЕСР форматами G7 та G20, – то інтеграція GPAI означає пряме поглинання автономної інституції. По-друге, вона засвідчує структурну закономірність, характерну для цієї сфери: окремі ініціативи з регулювання ШІ тяжіють до інтеграції в усталені міжнародні організації, що мають адміністративну інфраструктуру, стабільне фінансування та процедури, – потенціал, якого, як засвідчив досвід GPAI, бракує новоствореним інституціям. По-третє, інтеграція не усуває зазначеного вище обмеження репрезентативності: попри декларований намір розширення за рахунок країн, що розвиваються, інтегроване партнерство надалі не включає такі країни як Китай та Росію, що зберігає проблему легітимності його претензій на глобальну роль у регулюванні ШІ. Юридична природа актів інтегрованого GPAI залишається незмінною: це необов'язкові рекомендації та звіти, що мають переконувальний, а не примусовий характер. Інноваційним елементом є проєктний підхід GPAI, зокрема проєкт RAISE (Стратегія відповідального ШІ для довкілля), спрямований на застосування ШІ для вирішення екологічних проблем, що доповнює аналітичну діяльність ОЕСР практично орієнтованими ініціативами.

Окремий вектор міжрегіонального співробітництва сформували Саміти з безпеки ШІ, започатковані Декларацією Блетчлі (англ. – Bletchley Declaration), ухваленою 1–2 листопада 2023 року на першому такому саміті у Блетчлі-Парк (Велика Британія). Декларація, підписана 28 державами та Європейським Союзом, уперше на рівні політичного документа такого формату зафіксувала спільне визнання ризиків передових систем ШІ (англ. – frontier AI) та необхідність міжнародної співпраці в їх оцінюванні; її розвитком стали саміти в Сеулі (2024) та Парижі (2025). Україна долучилася до Декларації, що засвідчує її залученість до міжнародних безпекових ініціатив у сфері ШІ [285].

Критичний аналіз ролі міжрегіональних ініціатив у формуванні міжнародних стандартів застосування ШІ виявляє низку фундаментальних викликів та обмежень. Першим викликом є дихотомія між фрагментацією та консолідацією у глобальній архітектурі міжнародно-правового регулювання ШІ. Множина паралельних ініціатив різних міжнародних організацій створює ризики дублювання зусиль та конфліктів між актами. Проте поліцентричний підхід може забезпечувати більшу гнучкість, експериментування з різними моделями регулювання та адаптацію до особливостей різних юрисдикцій. Дослідники П. Сігон, М. Маас та Л. Кемп аргументують, що «деякі вчені вважають «поліцентричні» підходи до управління більш креативними та легітимними, ніж централізовано координовані режими», оскільки «поліцентричність дозволяє управлінським ініціативам починати справляти вплив на різних рівнях та уможлиблює експериментування з політиками та підходами» [148, с. 551]. Л. Шмітт підкреслює, що міжнародно-правове регулювання ШІ «є поліцентричним, оскільки має різні епістемічні спільноти та численні центри ухвалення рішень, кожен з яких діє з певним ступенем автономії. Воно є фрагментованим через суттєве перекриття у членстві різних акторів та тематиці, що розглядається цими ініціативами» [142, с. 311].

Другий виклик стосується вибору між «м'яким» і «твердим правом» як основними інструментами регулювання. Усі проаналізовані акти міжрегіонального рівня – Рекомендація ОЕСР, Принципи G20, Хіросімські принципи та Кодекс

поведінки G7, звіти та рекомендації GRAI – належать до категорії «м'якого права», що не створює юридично обов'язкових зобов'язань і не передбачає механізмів примусового виконання. Л. Шмітт зауважує, що «глобальна координація у цій сфері поки що не торкнулася юридично обов'язкових договорів. Цілком можливо, що уряди вирішили передати певну владу міжнародним організаціям лише доти, доки вони мають справу з досить абстрактними принципами чи м'яким управлінням, але відступили б або затягували, як тільки робота просунеться до більш регуляторного, твердого управління» [142, с. 311]. Ця закономірність – готовність держав до нормотворчої співпраці обернено пропорційна до юридичної обов'язковості передбачуваних зобов'язань – є характерною рисою міжрегіонального рівня регулювання ШІ.

Третім викликом є питання легітимності та репрезентативності. ОЕСР об'єднує переважно розвинені країни; G7 складається виключно з промислово розвинених демократій; GRAI не включає Китай та Росію. Навіть більш інклюзивний G20 не охоплює більшості держав – членів ООН. Хоча Рекомендація ОЕСР наразі має сорок сім прихильників, включно з деякими країнами, що розвиваються, фундаментальні питання щодо представництва глобального Півдня та специфічних потреб країн з різними рівнями технологічного розвитку залишаються актуальними. П. Сігон, М. Маас та Л. Кемп відзначають, що «країни, що розвиваються, не є належно представленими на більшості міжнародних зустрічей щодо ШІ» [148, с. 548], а така фрагментованість «може змушувати держави розпорошувати ресурси та фінансування між багатьма різними інституціями, обмежуючи спроможність менш забезпечених сторін до участі» [148, с. 549]. Геополітичний вимір створює додаткові складнощі для універсальності стандартів міжрегіонального рівня. Китай, який є другою за потужністю країною у розробці ШІ після США, має власну систему регуляторних підходів. Хоча Китай підтримав Принципи G20 зі ШІ, Л. Шмітт зауважує: «Поки Китай залишається осторонь більшості згаданих вище ініціатив, його роль в управлінні ШІ не можна недооцінювати. Уряд сигналізував готовність займатися глобальним управлінням як відповідальний актор, і конкретно щодо

етики ШІ зробив деякі кроки до примирення. Проте, він хотітиме інтерпретувати етику ШІ відповідно до свого власного культурного контексту та просувати ці погляди глобально» [142, с. 311]. Ця різноманітність інтерпретацій ставить під питання можливість досягнення справжньої конвергенції у регулюванні ШІ навіть за умови формальної згоди щодо базових принципів.

Четвертий виклик стосується співвідношення між державою та недержавними акторами. Хоча Рекомендація ОЕСР та Хіросімські принципи адресуються до різних категорій «акторів ШІ», механізми забезпечення дотримання стандартів недержавними акторами залишаються слабкими і покладаються переважно на добровільне прийняття та репутаційні стимули. Г. Робертс зазначає: «приватні інституції є ефективними у управлінні питаннями, пов'язаними з технологічною та транзакційною взаємопов'язаністю, оскільки це відповідає їхнім фінансовим інтересам. Проте, вони позбавлені стимулу значущим чином вирішувати питання, пов'язані з екологічними збитками чи безпекою ШІ, що свідчить про необхідність втручання публічного сектору» [150, с. 9]. За оцінкою Г. Робертса, «незважаючи на те, що ця увага є позитивною та надала імпульс раніше повільним зусиллям з глобального управління ШІ, є також підстави для песимізму. Існуючі ініціативи є загалом необов'язковими, позбавлені деталізації та здаються повторюваними. Через це вони мали обмежений вплив на зміну поведінки публічних та приватних акторів, що розробляють та розгортають ШІ» [150, с. 1].

П'ятий виклик пов'язаний зі спроможністю міжрегіональних інституцій вчасно адаптуватися до технологічних змін. Класична парадигма міжнародного нормотворення, орієнтована на створення стабільних довготривалих міжнародних договорів, за своєю природою не пристосована до регулювання технології, ландшафт якої може докорінно змінитися протягом одного-двох років, тоді як конвенційний процес зазвичай розтягується на десятиліття. Водночас саме на міжрегіональному рівні (і частково універсальному) сформувалася якісно нова модель нормотворення, спрямована на подолання цього «темпорального розриву».

Теоретичну рамку для осмислення цієї моделі надає доктрина інкременталізму у міжнародному правотворенні, що сформувалася у працях У. Хетавей [282] та С. Блок-Ліб і Т. Халлідєя [281]. С. Блок-Ліб та Т. Халлідєй у класичній роботі «Інкременталізми у глобальному правотворенні» (2007) на матеріалі діяльності ЮНСІТРАЛ виокремили три форми інкременталізму: «вертикальний інкременталізм має місце, коли міжнародні організації глибше занурюються у конкретну сферу через прогресивні раунди; горизонтальний інкременталізм можна спостерігати, коли міжнародні організації розширюють субстантивні межі діапазону тем ... через послідовні раунди; пірамідальний інкременталізм виникає, коли міжнародна організація свідомо створює свої норми, стоячи на плечах попередніх зусиль інших міжнародних організацій» [281, с. 854].

На основі проведеного аналізу пропонуємо визначати інституційну відповідь міжнародних організацій універсального та міжрегіонального рівнів на «темпоральний розрив» як «ітеративне нормотворення» – модель формування актів «м'якого права» (рекомендацій, стандартів, керівництв тощо), за якої один і той самий акт не розглядається як остаточна фіксація норм, а функціонує як «живий документ» (англ. – living document), що проходить повторювані цикли прийняття, апробації, оцінки ефективності та перегляду. Запропонована модель характеризується трьома сутнісними ознаками: нормативною незавершеністю як свідомим методологічним вибором, а не недоліком; суттєвим скороченням тривалості нормотворчих циклів – у низці випадків до одного-двох років; інкрементальністю змін при збереженні ідентичності акта, його архітектури та ціннісного ядра. Термін «ітеративне» свідомо запозичено з методології розробки програмного забезпечення, де ітерація означає повторюваний цикл створення, тестування та вдосконалення продукту з урахуванням результатів попередніх версій.

Запропонована концепція співвідноситься з вертикальним інкременталізмом С. Блок-Ліб та Т. Халлідєя як вид з родом, проте істотно відрізняється від нього за двома ключовими параметрами, що дозволяє розглядати її як самостійний теоретичний внесок у межах ширшої доктрини інкременталізму. По-перше,

класичний вертикальний інкременталізм описує рух між різними нормативними формами – від рекомендацій до модельних законів, від модельних законів до конвенцій, – тоді як ітеративне нормотворення передбачає збереження тієї самої нормативної форми (акта «м'якого права») при оновленні його змістового наповнення. По-друге, у моделі С. Блок-Ліб та Т. Халлідея інкрементальність впливає з внутрішньої логіки правотворчого процесу та обмеженості ресурсів міжнародних організацій, тоді як у запропонованій нами моделі вона є свідомою регулятивною відповіддю на зовнішній чинник – «темпоральний розрив» між швидкістю технологічних змін і повільністю нормативних процесів, який не є ціннісно нейтральним: адже, як зазначається в доктрині, регуляторна пауза дозволяє акумулювати шкоду до створення процедурних бар'єрів [283]. Подібним чином концепція ітеративного нормотворення відрізняється від теорії рекурсивності права Т. Халлідея та Б. Каррутерса [279, с. 1135] за природою рушійних механізмів (свідома інституційна стратегія проти непередбачуваного наслідку внутрішніх суперечностей) та предметним полем (новітні технології проти міжнародного економічного права).

У вітчизняній доктрині вживається концепція «регуляторної оперативності», що передбачає інструменти, які «вбудовують гнучкість на етапі нормопроекування: часові обмеження дії норми з обов'язковим переглядом, регуляторні пісочниці (обмежене та контрольоване випробовування нових рішень), поетапне та ітеративне нормотворення з чітким журналюванням змін і перехідними положеннями» [284, с. 334]. Термін «ітеративне нормотворення» вжито тут одинично – у переліку інструментів національного нормопроекування, без визначення, виокремлення сутнісних ознак чи доктринального позиціонування, – що засвідчує входження відповідної термінології у вітчизняний науковий обіг, водночас залишаючи відкритим завдання її концептуалізації. У цьому дослідженні таку концептуалізацію здійснено щодо міжнародно-правового регулювання ШІ: ітеративне нормотворення обґрунтовано як модель формування актів «м'якого права» міжнародними організаціями універсального та міжрегіонального рівнів – із дефініцією, трьома

сутнісними ознаками та позиціонуванням у межах доктрини інкременталізму, викладеними вище. Водночас запропонована модель резонує з ширшою методологічною траєкторією «права, що адаптується» (англ. – adaptive law), яка передбачає поетапність, експериментальність та швидку корекцію норм, і з моделлю «багаторівневого врядування ШІ» (англ. – layered governance), що комбінує базові правові гарантії з гнучкими технічними специфікаціями та галузевими стандартами [284, с. 348].

Парадигматичним прикладом ітеративного нормотворення є еволюція Рекомендації ОЕСР, що пройшла два цикли перегляду одного й того самого акта: у листопаді 2023 року оновлено визначення «системи ШІ» для відображення специфіки генеративного ШІ, а у травні 2024 року здійснено масштабнішу ревізію, що розширила предмет Рекомендації на питання приватності, інтелектуальної власності, безпеки та цілісності інформації [9]. На універсальному рівні аналогічну модель демонструє Резолюція 101 Всесвітньої асамблеї зі стандартизації електров'язку МСЕ, що періодично переглядається на конференціях ВАСЕ та у Нью-Делі 2024 року отримала оновлений мандат Сектору стандартизації електров'язку МСЕ-Т щодо ШІ [109]. У той же час, прикладом горизонтального інкременталізму – «логіки екстенсифікації дії, коли широта теми або сфери розширюється» [281, с. 900] – є Хіросімський процес G7, який протягом 2023–2024 рр. послідовно продукував Хіросімські керівні принципи, Міжнародний кодекс поведінки та Комплексну рамку політики, формуючи нормативний зміст через послідовне розширення тематичного охоплення взаємодоповнюючими документами. Прикладом пірамідального інкременталізму – коли міжнародна організація «свідомо створює свої норми, стоячи на плечах попередніх зусиль» інших [281, с. 854], – є Принципи G20 зі ШІ 2019 року, прямо запозичені з Рекомендації ОЕСР [10], та інституційна інтеграція GPAI з ОЕСР у 2024 році, коли первісна автономна архітектура Партнерства була перебудована через входження в інституційну структуру ОЕСР [277]. На універсальному рівні цю ж форму ілюструє

Глобальний цифровий договір, який свідомо інтегрує напрацювання ОЕСР, ЮНЕСКО, G7, G20 та інших форматів у єдину рамкову конструкцію [21]

Концепція ітеративного нормотворення має нормативне підґрунтя безпосередньо у самих актах: Рекомендація ЮНЕСКО закріплює принцип «адаптивного управління» з вимогою «вжити заходів для врахування зрушень у технологіях та появи нових груп зацікавлених сторін» [13, п. 47], ГЦД визнає необхідність «гнучкого, міждисциплінарного та адаптованого підходу з участю багатьох зацікавлених сторін» [21, п. 54], а Обсерваторія політики ОЕСД.АІ [8] функціонує як платформа моніторингу та зворотного зв'язку, що забезпечує інформаційну основу для подальших циклів перегляду. Ітеративне нормотворення стає можливим саме завдяки інструментам «м'якого права», які, на відміну від міжнародних договорів, не потребують тривалих процедур ратифікації: у цьому контексті «м'якість» перетворюється з обмеження на інституційну перевагу. Разом з тим, як прямо вказувала ще У. Хетавей, «небезпека цього підходу полягає в тому, що він може зупинитися на будь-якому етапі циклу» [282, с. 531], – такий характер змін може бути недостатнім для відповіді на технологічний зсув, а навіть своєчасно оновлені акти «м'якого права» залишаються неефективними за відсутності механізмів верифікації їх дотримання. Це вказує на необхідність доповнення ітеративного нормотворення іншими регуляторними інструментами, зокрема юридично обов'язковими актами регіонального рівня, аналіз яких є предметом наступного розділу.

У підсумку, роль міжрегіональних ініціатив у міжнародно-правовому регулюванні ШІ може бути охарактеризована як роль «м'яких консолідаторів» у поліцентричному режимному комплексі. ОЕСР виступає центральною ланкою цього кластеру завдяки своєму епістемічному авторитету, методологічній експертизі та інституційним зв'язкам. G7 забезпечує політичний імпульс та оперативну реакцію на нові технологічні виклики, як продемонстрував Хіросімський процес. G20 надає ширшу політичну легітимність узгодженим стандартам, хоча ціною значної абстрактності зобов'язань. GPAI, яке після інтеграції з ОЕСР у липні 2024 року

функціонує безпосередньо в її інституційній структурі, доповнює цю архітектуру експертним та проєктним компонентом, об'єднуючи фахівців з різних секторів для практичної імплементації принципів; при цьому сам факт інтеграції засвідчує тенденцію до консолідації раніше автономних ініціатив навколо усталених міжнародних організацій, що посилює центральну роль ОЕСР у міжрегіональному кластері. Усі ці інституції діють виключно в площині «м'якого права», не маючи ні формальної регуляторної влади, ні механізмів примусу, і використовують свій авторитет для створення спільних референтних рамок, полегшення діалогу та підтримки конвергенції національних підходів. Ця роль є одночасно цінною, бо у відсутності політичної можливості для створення універсального «твердого права» вона дозволяє досягати певного рівня міжнародної координації; та обмеженою, бо не забезпечує ні універсальності охоплення, ні ефективності виконання, ні справедливості результатів. Майбутнє міжнародно-правового регулювання ШІ потребуватиме поєднання різних інструментів та підходів, при якому «м'яке право» міжрегіональних ініціатив доповнюватиметься регіональними обов'язковими режимами, секторальними міжнародними договорами та новими формами багатостороннього регулювання.

Висновки до Розділу 2

Проведене дослідження стану і тенденцій міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях, що ґрунтується на системному аналізі 23 резолюцій ГА ООН за період 2015–2025 років, нормативних актів та ініціатив спеціалізованих установ системи ООН, а також актів та практик ОЕСР, G7, G20 та GPAI, дозволяє сформулювати такі узагальнюючі висновки.

На обох досліджуваних рівнях цілковито домінують інструменти «м'якого права»: жоден із проаналізованих актів не має юридично обов'язкової сили. Ці акти утворюють щільну мережу «м'якого права», яка характеризується, як зазначають Г. Робертс, М. Таддео та Л. Флоріді, не дихотомією «обов'язкове – необов'язкове», а

«континуумом за трьома вимірами – ступенем зобов'язання, точності та делегування» [150, с. 8]. У межах цього континууму нормативний потенціал актів суттєво різниться: від резолюцій ГА ООН, ухвалених консенсусом 193 держав, через Рекомендацію ЮНЕСКО з розгорнутою системою принципів та механізмами моніторингу імплементації, до аналітичних звітів ВОІВ та МОП, що виконують функцію попереднього дослідження. Принципи, закладені в основу цих актів, демонструють стійку конвергенцію навколо ядра спільних цінностей – людиноцентричного підходу, прозорості та пояснюваності, підзвітності та людського нагляду, безпеки та надійності, інклюзивності та недискримінації, захисту приватності та даних, – які з різним ступенем деталізації фіксуються у дванадцяти принципах резолюцій ГА ООН, Рекомендації ЮНЕСКО [13], шести принципах ВООЗ [113], п'яти принципах Рекомендації ОЕСР [9] та одинадцяти Хіросімських керівних принципах G7 [145]. Ця конвергенція не є випадковою: вона відображає перетворення «принципізму» (англ. – *principlism*) на «резонуючу основу для більшості різноманітних етичних рекомендацій, політичних рамок, правових орієнтирів, кодексів поведінки» [134, с. 4]. Разом з тим аксіологічний консенсус не означає повної тотожності підходів: ГА ООН наголошує на подоланні цифрового розриву та досягненні Цілей сталого розвитку (ЦСР); ЮНЕСКО додає вимоги щодо захисту культурного різноманіття та заборони соціального скорингу; ОЕСР підкреслює сприяння інноваціям; ВООЗ конкретизує загальні принципи через призму безпеки медичних систем ШІ.

Системі міжнародно-правового регулювання ШІ притаманна «градація нормативності», яка відображає функціональну спеціалізацію різних рівнів, установ та організацій і утворює своєрідні «концентричні кола»: у центрі – акти з найвищим ступенем політичної легітимності, але водночас із найвищою абстрактністю (резолюції ГА ООН, ухвалені консенсусом 193 держав); друге коло – акти спеціалізованих установ ООН, що конкретизують загальні принципи через секторальні механізми (Рекомендація ЮНЕСКО закріплює етичний фундамент, технічні стандарти МСЕ встановлюють вимоги інтеперабельності, керівництва

ВООЗ деталізують процедури для медичної сфери); третє коло – акти міжрегіональних ініціатив (Рекомендація ОЕСР, Принципи G20, Хіросімські принципи та Кодекс поведінки G7), що балансують між загальністю та конкретністю. Така градація є не лише описовою характеристикою чинної системи, а й її функціональною перевагою, оскільки дозволяє різним категоріям учасників відносин у сфері ІІІ знаходити відповідні точки дотику – від держав, які орієнтуються на резолюції ГА ООН, до суб'єктів приватного права – розробників і постачальників систем ІІІ, для яких релевантнішими є технічні стандарти МСЕ або рекомендації ОЕСР.

Міжнародно-правове регулювання ІІІ, що формується на універсальному та міжрегіональному рівнях, має поліцентричний та фрагментований характер. Взаємодія між рівнями є двоспрямованою: Рекомендація ОЕСР 2019 року концептуально вплинула на формулювання резолюцій ГА ООН 2024–2025 років, тоді як Рекомендація ЮНЕСКО 2021 року стала етичним фундаментом для ініціатив як інших спеціалізованих установ, так і міжрегіональних форматів. Розбіжності між рівнями мають інституційний та функціональний характер і виявляються у трьох вимірах: діапазоні охоплення, ступені деталізації приписів та механізмах імплементації. Жоден із цих механізмів не передбачає юридичних санкцій за недотримання встановлених стандартів, що створює асиметрію між легітимністю актів та їх практичною застосовністю: акти ГА ООН мають найвищу легітимність, але є найбільш абстрактними; акти ОЕСР та G7 є конкретнішими та придатнішими до безпосереднього застосування, але репрезентують лише частину міжнародної спільноти.

Досліджуваним рівням регулювання притаманні спільні стійкі та системні прогалини: жоден з рівнів не забезпечує адекватного регулювання таких критичних питань, як вплив технологій ІІІ на довкілля; відповідальність суб'єктів приватного права за наслідки функціонування систем ІІІ; виклики, пов'язані з генеративним ІІІ (дипфейки, «галюцинації», порушення авторських прав); проблематика загального штучного інтелекту (ЗІІІ, англ. – AGI) та довгострокових екзистенційних ризиків;

воєнне використання ШІ, яке, попри ухвалення ГА ООН серії резолюцій, не має дієвих імплементаційних механізмів на рівні спеціалізованих установ. Особливо гострим серед цих викликів залишається «темпоральний розрив» між швидкістю технологічного розвитку та повільністю нормотворчих процесів.

Розвиток міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях характеризується п'ятьма взаємопов'язаними тенденціями. Перша – інтенсифікація нормотворчої активності, емпірично підтверджена запропонованою чотиріперіодною хронологією від «Підготовчої фази» (2015–2022) через «Пробудження інтересу» (2023) і «Інституціоналізацію» (2024) до «Оперативної фази» (2025): 70% досліджуваних резолюцій ГА ООН (16 з 23) ухвалені протягом 2024–2025 років, з аналогічним прискоренням на рівні спеціалізованих установ та міжрегіональних ініціатив. Друга – перехід від фрагментарного до системного регулювання: Глобальний цифровий договір (ГЦД) став точкою конвергенції на універсальному рівні, а стандартизація триади «безпечні, захищені та надійні системи ШІ» і підходу на основі «життєвого циклу» систем ШІ свідчить про формування когерентної нормативної системи. Третя – інституціоналізація регулювання через створення постійних структур (Незалежна міжнародна наукова панель з ШІ, Глобальний діалог з управління ШІ, Обсерваторія політики ОЕСР у сфері ШІ, Обсерваторія МОП, інтеграція GPAI в інституційну структуру ОЕСР у 2024 році), що засвідчує перехід від ad hoc реагування до систематичної роботи. Четверта – розширення предмета та поглиблення деталізації регулювання: від загальних декларацій до детальної регламентації конкретних сфер застосування (воєнної, медичної, трудової) та специфічних технологічних викликів. П'ята – перехід до парадигми періодичного перегляду та оновлення актів «м'якого права» у відповідь на технологічні зміни.

Виявлена тенденція циклічного оновлення актів дістала теоретичне узагальнення в авторській концепції «ітеративного нормотворення». Під цією моделлю в дисертації розуміється такий порядок формування актів «м'якого права», за якого акт не розглядається як остаточна фіксація норм, а функціонує як «живий

документ» (англ. – living document), що проходить повторювані цикли прийняття, апробації, оцінки ефективності та перегляду. У межах доктрини інкременталізму в міжнародному правотворенні, що сформувалася у працях У. Хетавей [282] та С. Блок-Ліб і Т. Халлідея [281], запропонована концепція позиціонується як специфічна форма вертикального інкременталізму, адаптована до сфери регулювання технологій з непередбачуваною траєкторією розвитку, і відрізняється від класичної інкременталістської моделі за двома ключовими параметрами: збереженням тієї самої нормативної форми (акта «м'якого права») при оновленні його змістового наповнення (на відміну від класичного руху від рекомендацій до конвенцій) та обумовленістю зовнішнім чинником – «темпоральним розривом» між швидкістю технологічних змін і повільністю нормотворчих процесів. Концепція має нормативне підґрунтя безпосередньо у самих актах – у принципі «адаптивного управління» Рекомендації ЮНЕСКО [13, п. 47] та у визнанні Пактом про майбутнє необхідності «гнучкого, міждисциплінарного та адаптованого підходу з участю багатьох зацікавлених сторін» [21, п. 54]. Ітеративне нормотворення стає можливим саме завдяки використанню інструментів «м'якого права», які дозволяють оперативно коригувати нормативний зміст без необхідності повторного отримання формальної згоди держав, перетворюючи «м'якість» з обмеження на інституційну перевагу.

«М'яке право» у сфері регулювання ШІ виконує щонайменше чотири функції: кристалізації консенсусу (поступове узгодження принципів, термінології та концептуальних рамок формує спільне розуміння, що з часом може стати основою для обов'язкових норм); правового експерименту (тестування різних регуляторних підходів без ризиків передчасної фіксації обов'язкових норм у технологічній сфері, що швидко змінюється); легітимізації (формулювання принципів створює нормативні очікування, які впливають на поведінку учасників відносин через механізми соціального тиску та репутаційних стимулів); підготовки підґрунтя для юридично обов'язкового регулювання (принципи, апробовані як «м'яке право», поступово інкорпорується до юридично обов'язкових актів регіонального та

національного рівнів). Водночас саме лише «м'яке право» є недостатнім для ефективного регулювання технологій з масштабними та потенційно необоротними наслідками, до яких належить ШІ; його обмеженість зумовлює потребу в доповненні юридично обов'язковими інструментами, що формуються передусім на регіональному рівні, – такими як Регламент (ЄС) 2024/1689 (Акт про ШІ) [25] та Рамкова конвенція Ради Європи про штучний інтелект та права людини, демократію і верховенство права 2024 р. [24], – аналіз яких є предметом наступного розділу.

РОЗДІЛ 3. МІЖНАРОДНО-ПРАВОВЕ РЕГУЛЮВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ НА РЕГІОНАЛЬНОМУ РІВНІ

3.1. Формування правової бази Європейського Союзу у сфері регулювання штучного інтелекту

Дослідження міжнародно-правового регулювання ШІ на регіональному рівні потребує обґрунтування аналітичної послідовності, обраної в межах цього розділу, – а саме: аналізу формування правової бази ЄС, далі діяльності Ради Європи і, нарешті, порівняльної характеристики інших регіональних ініціатив. Така структура не є довільною, а визначається наступними взаємопов'язаними методологічними критеріями. По-перше, критерій регуляторної зрілості: саме ЄС є на сьогодні єдиним регіональним утворенням, що ухвалило юридично обов'язковий акт прямої дії – Регламент (ЄС) 2024/1689, – тоді як більшість інших регіональних організацій, перебувають на стадії формування відповідної нормативної бази. По-друге, критерій інституційної щільності: право ЄС характеризується найрозвиненішою та найбільш взаємопов'язаною системою регулювання ШІ, що охоплює базові регламенти та директиви, галузеві акти, а також розгалужену наглядову інституційну інфраструктуру. По-третє, критерій релевантності для України: в умовах євроінтеграційного курсу держави та її зобов'язань щодо поступового приведення національного законодавства у відповідність до *acquis* ЄС, саме його правова база має першочергове методологічне значення, що й обумовлює його пріоритетний розгляд. Ця послідовність забезпечує перехід від найбільш розробленої та конкретизованої системи регулювання – до порівняльного аналізу інших регіональних ініціатив.

Формування правової бази ЄС у сфері регулювання ШІ є складним багатоетапним правотворчим процесом, що розгортався протягом десятиріччя й увінчався прийняттям першого у світі комплексного юридично обов'язкового акта щодо ШІ – Регламенту (ЄС) 2024/1689 Європейського Парламенту та Ради про гармонізовані правила щодо штучного інтелекту (Акт про ШІ). На відміну від

універсального та міжрегіонального рівнів, на яких нормотворча діяльність розвивалася через акти «м'якого права», ЄС обрав якісно інший шлях: послідовного нарощування регуляторного потенціалу від документів рекомендаційного характеру до юридично обов'язкових актів. Цей підхід відображає специфіку інституційної архітектури ЄС, яка, на відміну від системи ООН, передбачає здійснення законодавчої функції Європейським Парламентом і Радою ЄС, що дозволяє ухвалювати регламенти прямої дії, безпосередньо обов'язкові для держав-членів та їхніх суб'єктів. Ретроспективний аналіз нормативного масиву дозволяє виокремити п'ять послідовних фаз у формуванні правової бази ЄС у сфері ШІ.

Першу фазу ознаменувала Резолюція Європейського парламенту від 16 лютого 2017 року з рекомендаціями Європейській Комісії щодо норм цивільного права у сфері робототехніки (2015/2103(INL)) [6], підготовлена Комітетом Європейського парламенту з правових питань. У Резолюції наголошувалося на необхідності врахування правових та етичних наслідків таких технологій, «не стримуючи при цьому інновацій» [6, п. В]. Цей принцип балансу між регулюванням та інноваціями став наскрізним для всієї подальшої правотворчої діяльності ЄС у сфері ШІ. Особливої уваги заслуговує та обставина, що Європейський парламент (далі – ЄП) вже у 2017 році запропонував створення спеціалізованого Європейського агентства з робототехніки та штучного інтелекту [6, п. 16] – ідея, яка хоча й не була реалізована у первісному вигляді, знайшла часткове втілення у створенні Європейського офісу зі штучного інтелекту (англ. – *AI Office*) у 2024 році, що свідчить про інституційну тяглість ідей, закладених ЄП. Крім того, Резолюція містила Хартію з робототехніки, що запропонувала кодекс етичної поведінки для інженерів та модельні ліцензії для дизайнерів і користувачів [6], демонструючи методологічну послідовність підходу ЄС, який із самого початку прагнув поєднати технічне регулювання з ціннісними орієнтирами.

Другу фазу формування правової бази – фазу стратегічного планування – відкрило Повідомлення Європейської комісії «Штучний інтелект для Європи» (COM(2018) 237 final) від 25 квітня 2018 року [152]. Цей програмний документ

уперше сформулював цілісну «Європейську ініціативу зі штучного інтелекту», що ґрунтувалася на трьох стратегічних стовпах: збільшенні державних та приватних інвестицій у ШІ, підготовці до соціально-економічних змін та забезпеченні належної етичної та правової бази [152]. Комісія прямо визнала проблему конкурентного відставання, зазначивши, що «без таких зусиль ЄС ризикує... залишитись споживачем рішень, розроблених в інших місцях» [152, с. 7], і встановила стратегічну мету досягнення обсягу інвестицій у ШІ щонайменше 20 мільярдів євро до кінця 2020 року з подальшим нарощуванням протягом десятиліття. Вже на цьому етапі виявляється характерна закономірність підходу ЄС: регуляторні ініціативи обґрунтовуються подвійним наративом – одночасно як засіб підвищення конкурентоспроможності і як інструмент захисту цінностей. Політичний імпульс стратегії надали висновки Європейської Ради, які ще у жовтні 2017 року закликали Комісію «представити європейський підхід до штучного інтелекту» [153, с. 8], а у червні 2018 року закріпили необхідність забезпечення «високоякісних даних» для розвитку ШІ [154, с. 3], заклавши підґрунтя для майбутніх Акту про управління даними [173] та Акту про дані [182]. На виконання цих політичних зобов'язань 7 грудня 2018 року Комісія опублікувала Координований план з розвитку та використання штучного інтелекту (COM(2018) 795 final) [155] – першу спільну стратегію Комісії та держав-членів, підтриману Радою у лютому 2019 року, яка закликала «держави-члени розробити до середини 2019 року національні стратегії або програми зі штучного інтелекту» [156, с. 9].

Третя фаза – формування етико-правового фундаменту – пов'язана з діяльністю Експертної групи високого рівня зі штучного інтелекту (AI HLEG), створеної Комісією у червні 2018 року і складеної з 52 експертів. Група підготувала два ключові документи: «Етичні настанови щодо надійного штучного інтелекту» у квітні 2019 року [47] та «Рекомендації щодо політики та інвестицій для надійного ШІ» у червні 2019 року [158]. Етичні настанови визначили три компоненти надійного ШІ (відповідність законодавству, дотримання етичних принципів та цінностей, технічну і соціальну надійність) та сформулювали сім ключових вимог:

людська діяльність та нагляд; технічна стійкість та безпека; конфіденційність та управління даними; прозорість; різноманітність, недискримінація та справедливість; соціальне та екологічне благополуччя; підзвітність. Ці сім вимог стали концептуальним ядром, що було згодом трансформовано в норми Акту про ШІ, хоча ступінь і характер такої трансформації виявилися неоднорідними. Структурно подібні принципи знайшли відображення у Рекомендації ОЕСР зі ШІ 2019 року [9] та Рекомендації ЮНЕСКО з етики ШІ 2021 року [13], що засвідчує формування міжінституційного консенсусу щодо базових принципів регулювання ШІ – не в останню чергу завдяки тому, що ряд експертів AI HLEG були залучені і до підготовки зазначених актів «м'якого права». Рекомендації щодо політики та інвестицій для надійного ШІ [158] запропонували 33 конкретні рекомендації, згруповані за чотирма напрямками (люди та суспільство; дослідження та промисловість; державний сектор; регулятивне середовище), у тому числі здійснення оцінки впливу на основоположні права перед впровадженням систем ШІ у чутливих сферах та створення загальноєвропейського механізму сертифікації. Обидві ці рекомендації знайшли часткове відображення в Акті про ШІ: ст. 27 передбачає обов'язок проведення оцінки впливу на основоположні права (FRIA) перед розгортанням систем ШІ високого ризику – хоча, як буде показано нижче, лише для обмежених категорій розгортачів та з винятками, що суттєво обмежили практичну реалізацію цих механізмів порівняно з початковим замислом.

Четверта фаза – перехід від етичних орієнтирів до юридично обов'язкового регулювання – розпочалася публікацією Білої книги «Штучний інтелект – Європейський підхід до досконалості та довіри» (COM(2020) 65 final) у лютому 2020 року [63]. Біла книга вперше на рівні програмного документа Комісії запропонувала ризикоорієнтований підхід до регулювання, визначивши два кумулятивних критерії для віднесення системи ШІ до високоризикових: по-перше, використання у секторах, де ризики є найбільш імовірними, та по-друге, використання таким чином, що значні ризики, ймовірно, виникнуть [63, с. 17]. Цей підхід мав значний вплив на архітектуру Акту про ШІ. Біла книга також визнала, що для систем ШІ з

неприйнятним ризиком можуть бути доцільними абсолютні заборони. Зокрема, було зазначено, що використання застосувань ШІ з метою дистанційної біометричної ідентифікації та інших інтрузивних технологій спостереження завжди вважатиметься високоризиковим [63, с. 18]. Ці попередні категоризації значною мірою визначили зміст ст. 5 Акту про ШІ щодо заборонених практик та Додатку III щодо високоризикових систем.

У квітні 2021 року Європейська комісія оприлюднила пропозицію Регламенту, що встановлює гармонізовані правила щодо штучного інтелекту (COM(2021) 206 final) [26], одночасно з переглянутим Координованим планом зі ШІ (COM(2021) 205 final) [159]. Пропозиція ґрунтувалася на ст. 114 ДФЄС, що регулює функціонування внутрішнього ринку. Цей вибір є принципово значущим для розуміння як можливостей, так і обмежень Акту: обираючи ст. 114 ДФЄС, Комісія позиціонувала регулювання ШІ передусім як інструмент гармонізації ринкових умов, а не як самостійний правозахисний механізм. Таким чином, Акт про ШІ був інтегрований у нормативну логіку єдиного цифрового ринку ЄС (англ. – EU Digital Single Market), у межах якого гармонізація технічних та безпекових вимог слугує усуненню бар'єрів для вільного обігу цифрових продуктів і послуг між державами-членами. Тому правозахисний вимір Акту реалізується опосередковано – через встановлення умов доступу до єдиного ринку лише для систем ШІ, що відповідають визначеним регуляторним вимогам, зокрема у сфері безпеки та прозорості. Процес прийняття Акту тривав понад два з половиною роки і виявився надзвичайно складним – зокрема, через появу генеративного ШІ, що змусило законодавців суттєво переглянути його початкову структуру. Дослідники ШІ Дж. Чун, К. Шредер де Вітт та К. Елкінс зазначають, що «випуск ChatGPT у листопаді 2022 року та його стрімке освоєння мільйонами споживачів по всьому світу застали європейських законодавців зненацька і призвели до суттєвих коригувань підходу Акту до регулювання ШІ» [194, с. 4]. Особливо гострою стала дискусія щодо регулювання моделей ШІ загального призначення (GPAI): Німеччина, Франція та Італія виступили проти жорсткого регулювання базових моделей, мотивуючи це

необхідністю підтримки європейських розробників. Ці переговори, що ледь не призвели до провалу прийняття Акту, завершилися досягненням домовленості 9 грудня 2023 року.

Регламент (ЄС) 2024/1689 Європейського парламенту і Ради від 13 червня 2024 року [25] набрав чинності 1 серпня 2024 року з поетапним графіком застосування: загальні положення та заборони почали застосовуватися з 2 лютого 2025 року; правила щодо управління та моделей GPAI – з 2 серпня 2025 року; загальне застосування, включаючи правила для високоризикових систем та зобов'язання щодо прозорості, – з 2 серпня 2026 року; а правила для високоризикових систем ШІ, що підпадають під чинне галузеве законодавство про безпеку продукції, – з 2 серпня 2027 року. Такий поетапний графік забезпечує суб'єктам достатній адаптаційний період і є безпрецедентним за рівнем деталізації серед усіх міжнародно-правових актів у сфері ШІ. Для полегшення переходу до нового регуляторного режиму Комісія запровадила «Пакт зі ШІ» (AI Pact) [161] – добровільну ініціативу, що запрошує розробників та операторів систем ШІ завчасно впроваджувати ключові зобов'язання Акту ще до набрання чинності відповідними положеннями. Цей інструмент «м'якого права» у перехідний період засвідчує прагнення Комісії забезпечити поступову адаптацію індустрії до нових вимог, уникаючи регуляторного шоку.

З огляду на значний обсяг Акту про ШІ, який налічує 113 статей та 13 додатків, подальший аналіз зосереджується на тих нормативних положеннях, які є найбільш релевантними для предмета цього дослідження: визначенні системи ШІ (ст. 3), заборонених практиках та класифікації ризиків (ст. 5, 6), вимогах до систем високого ризику (ст. 9–15), зобов'язаннях прозорості та підзвітності (ст. 16–27, 50), регулюванні моделей загального призначення (ст. 51–56), механізмі регуляторних «пісочниць» (ст. 57–59), інституційній архітектурі (ст. 65–68, 88), постринковому моніторингу та праві на оскарження (ст. 72, 85, 86) та санкційному механізмі (ст. 99). Такий підхід дозволяє виявити ключові закономірності регуляторної моделі ЄС, не перетворюючи аналіз на постатейний коментар до Акту.

Принципове значення для сфери застосування Акту має визначення «системи штучного інтелекту», закріплене у ст. 3(1). Відповідно до фінальної редакції, система ШІ визначається як «машинна система, спроектована для функціонування з різним ступенем автономії, яка може демонструвати адаптивність після розгортання і яка для явних або неявних цілей, на підставі отриманих вхідних даних, генерує результати – зокрема прогнози, рекомендації або рішення, – що можуть впливати на фізичне або віртуальне середовище» [25]. Це визначення суттєво відрізняється від запропонованого Комісією у пропозиції 2021 року (COM(2021) 206 final), яке було значно ширшим і ґрунтувалося на переліку конкретних технічних методів, зазначених у Додатку I (машинне навчання, логічні та статистичні підходи). Як зазначає Дубняк, «Акт про ШІ декларує технологічно нейтральне визначення «систем штучного інтелекту», однак не надає дефініції «штучний інтелект» [197, с. 51]. Еволюція визначення від широкого переліку технік до функціонального підходу відображає важливий компроміс законодавчого процесу, спрямований на досягнення балансу між регулюванням та запобіганням надмірній регуляторній обтяжливості.

Обрання парадигми «рамка безпеки продукції» (англ. - *product safety framework*) як основи регулювання ШІ є, ймовірно, найважливішим рішенням, що визначило всю структуру Акту. Регламент прямо базується на принципах Нової законодавчої рамки (англ. - *New Legislative Framework*), що передбачає систему гармонізованих стандартів, оцінки відповідності, маркування та ринкового нагляду. Проте ця аналогія з регулюванням продуктової безпеки є водночас і силою, і слабкістю Акту. Дослідники Дж. Чун, К. Шредер де Вітт та К. Елкінс зауважують, що «Акт спочатку був побудований у рамці безпеки продукції, але потім поєднаний із порядком денним щодо основоположних прав на вимогу Європейського парламенту та всупереч тиску з боку Європейської комісії» [194, с. 3]. М. Алмада та А. Раду розвивають цю думку, вказуючи, що «рамка безпеки продукції» не покриває всіх аспектів захисту прав людини і суспільних інтересів: «Захист основних прав через рамку безпеки продукції, передбачену Актом про ШІ, може бути

неефективним. Акт може не реагувати на проблеми, які неможливо легко сформулювати в рамках кількісного ризику, і він може не враховувати всі аспекти ризиків, які він виявляє» [193, с. 652]. Цей висновок є принципово важливим з точки зору міжнародного права прав людини: юридично обов'язковий інструмент, побудований переважно на основі гармонізації внутрішнього ринку, має іманентні обмеження у сфері захисту основоположних прав.

Центральним структурним елементом Акту про ШІ є ризикоорієнтований підхід, що класифікує системи ШІ на чотири категорії ризику: неприйнятний ризик (заборонені практики згідно зі ст. 5), високий ризик (Глава III), обмежений ризик (зобов'язання щодо прозорості за ст. 50) та мінімальний ризик (без спеціального регулювання). Ст. 5 Акту встановлює вичерпний перелік заборонених практик ШІ, що включає системи з підсвідомими, маніпулятивними чи обманними техніками; системи, які експлуатують вразливості осіб через вік, інвалідність або соціально-економічну ситуацію; системи соціального скорингу; системи біометричної категоризації за чутливими ознаками; а також системи біометричної ідентифікації в реальному часі у загальнодоступних місцях для цілей правоохоронної діяльності, за окремими суворо обмеженими винятками. С. Вахтер критично оцінює саме збереження цих винятків, зазначаючи, що «дистанційна біометрична ідентифікація має жахливий рівень точності, поліцейські системи для прогнозування можуть генерувати расистські та сексистські результати, а програмне забезпечення для розпізнавання емоцій майже не має здатності об'єктивно вимірювати реакції» [192, с. 680]. Як обґрунтовано в попередніх авторських дослідженнях, такі винятки для правоохоронних та міграційних органів створюють «зони без цифрових прав» у найбільш чутливих сферах застосування систем ШІ [172, с. 25]. Попри ці вади, сама модель класифікації за рівнем ризику залишається продуктивною: ризикоорієнтований підхід має «значний потенціал для подальшої адаптації на міжнародному рівні, оскільки дозволяє забезпечити баланс між захистом прав людини та стимулюванням інновацій» [243, с. 228]. Водночас його поширення

потребує уніфікації класифікаційних критеріїв задля формування узгодженого міжнародно-правового регулювання ШІ.

Категорія високоризикових систем ШІ, визначена у ст. 6 та Додатку III Регламенту, становить основне тіло регуляторної рамки Акту. Додаток III визначає вісім сфер високого ризику: біометрика, критична інфраструктура, освіта та професійна підготовка, зайнятість та управління працівниками, доступ до основних приватних і публічних послуг, правоохоронна діяльність, міграція та управління кордонами, відправлення правосуддя та демократичні процеси [25]. Українські науковці Н. Горобець та С. Науменко відзначають, що Акт «відносить судову систему до високоризикованих сфер застосування ШІ та висуває обов'язкові вимоги до технологій ШІ в цій сфері» [200, с. 17], що має безпосереднє значення для імплементації цих вимог у державах-членах та країнах-кандидатах. Німецька дослідниця І. Куше звертає увагу на принципово важливий момент: «класифікація високого ризику визначається не певним типом технічної оцінки ризиків, а результатом політичних суджень, які імпліцитно пов'язані з цінностями» [195, с. 8]. Цей парадокс є фундаментальним для розуміння всієї регуляторної архітектури Акту: ризикоорієнтований підхід, що позиціонується як об'єктивний та технічний, насправді є глибоко ціннісно навантаженим і політично зумовленим. Ще більш проблематичною є запроваджена у ст. 6(3) Акту можливість вилучення системи ШІ з категорії високоризикових навіть за використання у сферах, зазначених у Додатку III. С. Вахтер характеризує цю лазівку, засновану на самооцінці, як таку, що підриває основну мету регулювання: «постачальники не лише активно залучені до написання стандартів, яких вони мають дотримуватися, але й мають оцінювати самостійно, чи відповідають вони цим стандартам» [192, с. 692].

Система оцінки відповідності, встановлена ст. 43 Акту, є ключовим механізмом забезпечення ефективності ризикоорієнтованого підходу. Для високоризикових систем ШІ, що використовуються у сфері біометрики (п. 1 Додатку III), Регламент передбачає два альтернативних шляхи: внутрішній контроль або оцінку за участю нотифікованого органу; натомість для решти високоризикових

систем постачальники застосовують процедуру оцінки відповідності на основі лише внутрішнього контролю. Тож переважна більшість високоризикових систем ШІ – від критичної інфраструктури до правоохоронної діяльності та правосуддя – підлягає самооцінці постачальника без незалежної верифікації. Водночас, як зазначає Дж. Шуетт, система управління ризиками за ст. 9 Акту складається з двох компонентів – процесу управління ризиками (параграфи 2–4) та процедур тестування (параграфи 5–7), проте гармонізовані стандарти управління ризиками ШІ наразі відсутні, а їх створення CEN/CENELEC може тривати кілька років [189, с. 373]. К. Новеллі, Ф. Казоларі, А. Ротоло, М. Таддео та Л. Флоріді підсилюють цю критику, зазначаючи, що Акт «не має детальної методології оцінки ризиків для ідентифікації ризиків у конкретних ситуаціях, спираючись на статичне уявлення про ризик ШІ» [190, с. 2]. Це створює структурний розрив між амбітністю ризикоорієнтованого підходу та інструментальними засобами його реалізації, а вибір на користь самооцінки контрастує з підходом інших сфер продуктової безпеки ЄС: Регламент (ЄС) 2023/1230 про машинне обладнання [181] вимагає оцінки відповідності третьою стороною для всіх категорій машинного обладнання підвищеного ризику, а Регламент (ЄС) 2024/2847 про кібербезпеку продуктів з цифровими елементами [186] – залучення нотифікованих органів, що засвідчує порушення системної когерентності права ЄС.

Окрему увагу слід приділити принципам прозорості та підзвітності. Прозорість тлумачиться Актом через забезпечення «належної відстежуваності та пояснювальності» систем ШІ та інформування користувачів про взаємодію з ними та їхні права [25, с. 8], а підзвітність – через «запровадження механізмів для забезпечення відповідальності за системи ШІ та їх результати, як до, так і після їх розробки, розгортання та використання» [47, с. 19]. Прозорість реалізується через вимогу пояснювальності (ст. 13), технічну документацію (ст. 11), обов'язок інформування про взаємодію з ШІ (ст. 50) та реєстрацію систем високого ризику в Базі даних ЄС (ст. 71); підзвітність забезпечується через обов'язки учасників створення та постачання систем (ст. 16–26), процедури оцінки відповідності (ст. 40

–43), інституційний нагляд (ст. 56, 75), постринковий моніторинг (ст. 72), право на оскарження (ст. 85, 86) та юридичну відповідальність (ст. 99). Водночас реалізація цих принципів стикається з трьома суттєвими структурними перешкодами. Першою є вплив лобістських зусиль технологічних корпорацій на зміст кінцевого тексту Акту: як зазначає С. Вахтер, «надмірна залежність від саморегулювання, самосертифікації, слабких процедур нагляду та розслідування, а також далекосяжні винятки як для державного, так і для приватного секторів є продуктом цього лобіювання» [192, с. 674], – конкретним проявом чого стало запровадження механізму самокласифікації за ст. 6(3). Другою перешкодою є широкі винятки для правоохоронних та міграційних органів, які відповідно до ст. 49(4) Акту реєструють лише обмежену інформацію у непублічно доступній частині бази даних ЄС [25], що виключає можливість публічного контролю з боку громадянського суспільства та правозахисних організацій і фактично розширює описану вище «зону без цифрових прав» на сферу реєстраційних процедур. Третьою перешкодою є інституційна незрілість наглядових органів держав-членів ЄС: брак технічної експертизи, недостатність фінансових і людських ресурсів та відсутність усталеної практики правозастосування у сфері ШІ. Сукупність цих факторів створює ризик того, що Акт про ШІ, попри амбіційність своїх положень, не зможе забезпечити ефективний захист прав осіб, постраждалих від використання систем ШІ.

Щодо регулювання систем ШІ загального призначення (GPAI), яке було включено до Акту лише на пізніх етапах законодавчого процесу, Глава V Регламенту (ст. 51–56) встановлює дворівневу систему: загальні зобов'язання для всіх постачальників моделей GPAI та додаткові вимоги для моделей із системним ризиком. Ст. 51(2) визначає презумпцію наявності системного ризику, коли кумулятивний обсяг обчислень, використаних для навчання моделі, перевищує 10^{25} операцій з плаваючою комою [25]. С. Вахтер критично оцінює цю дворівневу модель, зазначаючи, що багато системних ризиків виникають у всіх моделях GPAI, незалежно від їхнього розміру чи обсягу обчислень: «дезінформація, галюцинації, упередженість, витіснення робочих місць, проблеми захисту даних та шкідливі

наслідки трапляються у менших і менш потужних системах» [192, с. 688]. Важливим механізмом конкретизації зобов'язань постачальників моделей GРАІ є Кодекс практики, розроблений за сприяння Європейського офісу зі ШІ відповідно до ст. 56 Акту про ШІ і структурований за трьома розділами – Прозорість [166], Авторське право [167] та Безпека і захищеність [168], – який виконує функцію «керівного документа для демонстрації відповідності зобов'язанням, передбаченим ст. 53 та 55 Акту про ШІ» [166]; розділ про авторське право, зокрема, конкретизує зобов'язання щодо дотримання застереження прав правовласниками відповідно до ст. 4(3) Директиви (ЄС) 2019/790 [171]. Практичним інструментом реалізації зобов'язань щодо прозорості є Форма документування моделей [169], що стандартизує обсяг та структуру інформації, яку постачальники GРАІ зобов'язані надавати відповідно до ст. 53(1)(a)–(d) Акту, і є першою спробою ЄС перевести абстрактні вимоги прозорості у конкретний, уніфікований формат звітності.

Проблема делегування стандартизації заслуговує на окрему увагу. Ст. 40(1) Регламенту встановлює, що системи ШІ високого ризику, які відповідають гармонізованим стандартам, «презюмуються такими, що відповідають вимогам, встановленим у статтях 9–15 Акту про ШІ» [25]. Комісія підготувала запит на стандартизацію до Європейського комітету зі стандартизації (СЕН) та Європейського комітету з електротехнічної стандартизації (СЕНЕЛЕС), які спільно працюють у рамках Технічного комітету JTC 21 [162]. Як зазначається у запиті на стандартизацію, «гармонізовані стандарти та результати стандартизації, прийняті на підтримку Акта про ШІ, мають забезпечувати високий рівень захисту здоров'я, безпеки та основоположних прав усіх осіб у Союзі» [163, с. 2]. Однак органи стандартизації не мають досвіду у сфері основоположних прав, і тому процес встановлення стандартів, ймовірно, буде імпліцитно надавати пріоритет одним цінностям над іншими. Формально ризикоорієнтований підхід ЄС фактично перетворюється на «стандартоорієнтований» підхід, тобто такий, де ключові рішення про баланс цінностей приймаються технічними органами з обмеженою компетенцією у сфері прав людини.

Інституційна архітектура Акту базується на складній багаторівневій системі. Рішенням Комісії від 24 січня 2024 року (C/2024/1459) було створено Європейський офіс зі штучного інтелекту, на який покладено функції підтримки безпечного розвитку та використання технологій ШІ в Союзі [160]. Офіс здійснює нагляд за дотриманням правил щодо моделей GPAI, координує діяльність національних компетентних органів та уповноважений здійснювати виключний нагляд відповідно до ст. 88 Акту, що закріплює за Комісією «виключні повноваження щодо нагляду та правозастосування» таких систем ШІ [25]. Координаційним органом на рівні держав-членів є Рада з ШІ, створена відповідно до ст. 65 Акту, яка забезпечує співпрацю між національними наглядовими органами та видає рекомендації щодо застосування Акту [25]. Її композиційний підхід – виключне представництво національних наглядових органів – компенсується створенням Консультативного форуму (ст. 67), що формується зі «збалансованого представництва зацікавлених сторін, включаючи промисловість, стартапи, МСП, громадянське суспільство та академічну спільноту» [25], та Наукової групи незалежних експертів (ст. 68), створеної для «підтримки правозастосувальної діяльності відповідно до цього Регламенту» [25]. Ця триланкова структура – Рада, Консультативний форум та Наукова група – є інституціоналізацією на регіональному рівні моделі управління за участю багатьох зацікавлених сторін у межах юридично обов'язкової регуляторної рамки, що принципово відрізняє її від аналогічних форматів на міжрегіональному рівні (Робоча група ОЕСР з управління ШІ та Мережа експертів ОЕСР зі штучного інтелекту [136]). Проте ні Консультативний форум, ні Наукова група не мають повноважень ухвалювати обов'язкові рішення, що відтворює на регіональному рівні структурну проблему дієвості зовнішнього контролю.

Окремим інноваційним елементом Акту є регуляторні «пісочниці» для ШІ, передбачені ст. 57–59. Необхідність їх запровадження обґрунтовується тим, що ШІ «є технологічним сімейством, що швидко розвивається, і потребує регуляторного нагляду та безпечного контрольованого простору для експериментування, водночас забезпечуючи відповідальну інновацію та інтеграцію належних гарантій і заходів

зниження ризиків» [25]. Кожна держава-член зобов'язана створити принаймні одну «пісочницю» до 2 серпня 2026 року – контрольоване середовище, що «сприяє інноваціям та полегшує розробку, навчання, тестування та валідацію інноваційних систем ШІ протягом обмеженого часу до їх розміщення на ринку або введення в експлуатацію відповідно до конкретного плану «пісочниці», узгодженого між постачальниками або потенційними постачальниками та компетентним органом» [25]. Перетворення «пісочниць» із факультативного інструменту політики у юридичний обов'язок держав-членів принципово відрізняє підхід ЄС від рекомендаційного характеру аналогічних механізмів у практиці ОЕСР та від пропозиції Білої книги системи ООН «забезпечити функціонування пісочниць для сприяння розробці міжнародно гармонізованих підходів до оцінки ризиків штучного інтелекту та зусиль з моніторингу» [128, с. 6], яка залишається на рівні побажання. Участь у «пісочниці» не звільняє учасників від дотримання чинного законодавства – вона лише створює умови для тестування з можливістю отримання регуляторних рекомендацій.

Загалом інституційна архітектура Акту про ШІ демонструє якісний стрибок порівняно з інституційними механізмами універсального та міжрегіонального рівнів: якщо Обсерваторія ОЕСР інформує, Робоча група AIGO координує, а Методологія оцінки готовності ЮНЕСКО [106] діагностує, то система Акту про ШІ регулює, примушує та санкціонує. Ключовим інструментом примусу є санкційний механізм із тривірневою градацією штрафів: до 35 мільйонів євро або 7% глобального річного обороту за заборонені практики (ст. 5), до 15 мільйонів євро або 3% обороту за порушення інших положень та до 7,5 мільйонів євро або 1% обороту за надання недостовірної інформації. Цей режим є більш диференційованим, ніж GDPR, і перевершує його за ступенем суворості щодо найбільш небезпечних порушень, що забезпечує ефективний примус там, де ризики для прав людини є найсерйознішими.

Комплексний аналіз правової бази ЄС у сфері ШІ вимагає розуміння того, що Акт про ШІ функціонує не як ізольований нормативний акт, а як центральний

елемент ширшої нормативної екосистеми, в якій різні регламенти та директиви перебувають у відносинах функціональної взаємодоповнюваності. Найбільш значущим є перетин Акту з Загальним регламентом про захист даних (GDPR) [170]: ст. 22 GDPR забороняє рішення, що базуються виключно на автоматизованій обробці та мають правові або аналогічно суттєві наслідки для особи, а ст. 14 Акту через вимоги людського нагляду забезпечує реалізацію цього права; ст. 27(4) Акту встановлює, що оцінка впливу на основоположні права (FRIA) доповнює оцінку впливу на захист даних (DPIA) за ст. 35 GDPR, охоплюючи ширший спектр впливу на права. Водночас, як зазначають Г. Ласкос та П. де Герт, поняття «високого ризику» в Акті про ШІ (ст. 6) та в GDPR (ст. 35(1)) не є тотожними, що створює додаткову складність для правозастосування [191]. Другим ключовим перетином є взаємодія з Актом про цифрові послуги (DSA) [175]: С. Вахтер зазначає, що механізми зовнішнього аудиту та доступу дослідників до даних великих платформ (ст. 37, 40 DSA) могли б бути перенесені на моделі GPAI із системним ризиком, проте відповідні положення Акту обмежуються переважно внутрішньою оцінкою постачальників [192, с. 688].

До цієї нормативної екосистеми також належать Акт про цифрові ринки [174], що регулює діяльність великих технологічних компаній – ключових розповсюджувачів систем ШІ; Акт про управління даними [173] та Акт про дані [182], що забезпечують правові умови для обміну даними для тренування систем ШІ; Директива NIS2 [176] і Регламент DORA [177] із вимогами кібербезпеки для критичної інфраструктури та цифрової операційної стійкості фінансового сектору; Регламент eIDAS 2.0 [183], що створює інфраструктуру цифрової ідентифікації; Регламент EHDS [187], що регулює вторинне використання медичних даних для навчання систем ШІ; Загальний регламент про безпеку продукції (GPSR) [180], який встановлює горизонтальну рамку безпеки для продуктів з інтегрованими компонентами ШІ; а також Директива (ЄС) 2024/2853 про відповідальність за дефектну продукцію [185], що вперше включає програмне забезпечення, у тому числі системи ШІ, до категорії «продуктів», доповнюючи превентивний підхід Акту

компенсаторним механізмом захисту. Сукупно ці акти формують нормативний каркас єдиного цифрового ринку ЄС, у межах якого регулювання ШІ постає не як відокремлений режим, а як складова правової бази ЄС у цифровій сфері.

Особливе значення має Директива (ЄС) 2024/2831 про покращення умов праці на платформах [184] – перший секторальний акт ЄС, що конкретизує загальні вимоги Акту про ШІ у контексті трудових відносин через спеціальний режим «алгоритмічного управління»: вона встановлює обмеження на автоматизовані системи моніторингу та прийняття рішень (ст. 7), вимагає людського нагляду (ст. 10) та гарантує працівникам право на пояснення автоматизованих рішень (ст. 11), а її ст. 7(1)(а) та 7(1)(е) забороняють обробку даних щодо емоційного стану та біометричних даних працівників, що кореспондує із забороненими практиками за ст. 5 Акту про ШІ. Така структурна координація між горизонтальним актом та галузевим законодавством демонструє тенденцію до «вертикалізації» регулювання ШІ. Описана щільність регуляторної мережі є безпрецедентною у глобальному масштабі та принципово відрізняє підхід ЄС від «м'якого» регулювання універсального та міжрегіонального рівнів; водночас вона породжує проблему когерентності, оскільки неузгодженість понятійного апарату, різні моделі оцінки ризиків та нерівномірна глибина регуляторних вимог створюють ризик фрагментації та ускладнюють правозастосування.

П'ята, сучасна фаза формування правової бази ЄС характеризується розгортанням масштабних стратегічних ініціатив щодо практичної імплементації Акту та зміцнення конкурентоспроможності ЄС. На *AI Action Summit* у Парижі 11 лютого 2025 року було анонсовано ініціативу *InvestAI* – програму мобілізації публічних та приватних інвестицій у ШІ загальним орієнтовним обсягом 200 мільярдів євро, призначеного, зокрема, для фінансування чотирьох «гігафабрик» з ШІ на території ЄС [164]. У квітні 2025 року Комісія представила План дій «Континент ШІ» (COM(2025) 165 final), який визначає стратегічні пріоритети у сферах інфраструктури обчислень, доступу до даних, кадрового потенціалу та єдиного ринку з чіткими правилами [165, с. 3]. Цей стратегічний курс вписується у

ширшу цифрову стратегію Союзу, спрямовану на завершення розбудови єдиного цифрового ринку ЄС, – Програму політики цифрового десятиліття до 2030 [178] та Європейську декларацію цифрових прав і принципів для цифрового десятиліття [179], – формуючи єдину рамку, в якій регулювання ІІ є одним із компонентів комплексної цифрової трансформації. Масштаб запланованих інвестицій – 200 мільярдів євро – у десять разів перевищує стратегічну мету 2018 року, що засвідчує перехід від переважно регуляторного фокусу (2017–2024) до комбінації регулювання та активної промислової політики, обумовленої посиленням геополітичної мотивації ЄС у глобальній гонці за ІІ.

Міжнародний вимір Акту про ІІ проявляється через феномен «Брюссельського ефекту», теоретично обґрунтований А. Бредфорд [202]: ст. 2 Акту закріплює екстратериторіальний механізм, що поширює його дію на будь-яку систему ІІ, результати якої використовуються на території ЄС, навіть якщо її постачальники розташовані у третій країні [25]. М. Алмада та А. Раду зазначають, що «Акт про ІІ від самого початку був розроблений з урахуванням його глобальних наслідків» [193, с. 653], – проте такий ефект водночас може обмежити простір для розробки більш ефективних правозахисних механізмів в інших регіонах та на універсальному рівні. Найяскравішим проявом цього ефекту є хронологічний та змістовний зв'язок Акту з Рамковою конвенцією Ради Європи про штучний інтелект, прийнятою 17 травня 2024 року [24] – лише через два місяці після остаточного голосування Парламенту щодо Акту 13 березня 2024 року [27]: ще у листопаді 2022 року Рада ЄС уповноважила Комісію вести переговори з метою забезпечення сумісності позицій ЄС з Радою Європи в напрямку ризикоорієнтованого підходу, «повністю сумісного з Актом про ІІ», та визнання «важливої ролі технічних стандартів та механізмів сертифікації» [193, с. 661].

У порівняльній перспективі з універсальним та міжрегіональним рівнями Акт про ІІ принципово відрізняється від ключових актів «м'якого права» – Рекомендації ОЕСР [9], Рекомендації ЮНЕСКО [13] та резолюцій ГА ООН – не за ціннісним фундаментом, а за юридичною природою: якщо акти універсального та

міжрегіонального рівнів забезпечують глобальний ціннісний консенсус (зокрема щодо принципів прозорості, підзвітності, людського контролю), то Акт ЄС надає найбільш деталізовану та юридично обов'язкову регуляторну рамку, операціоналізуючи ці принципи через конкретні зобов'язання та механізми примусу.

Виявлені закономірності та фази формування правової бази ЄС у сфері регулювання ІІІ мають безпосереднє значення для України, яка перебуває на початковому етапі побудови власного правового регулювання ІІІ. Як зазначає О.Г. Кожухар, «для України розуміння досвіду ЄС є важливим... На шляху до власного підходу до ІІІ Україні доцільно орієнтуватися на стандарти ЄС» [196, с. 65], що зумовлено європейським інтеграційним курсом держави та її зобов'язаннями щодо наближення законодавства до *acquis* ЄС відповідно до Угоди про асоціацію. Як підкреслює О.В. Олійник, Україна, «перебуваючи на етапі нормативного становлення у цій сфері, зіштовхується з низкою викликів, що зумовлює необхідність осмислення міжнародного досвіду, зокрема підходів Європейського Союзу та США» [201, с. 3]. Ю. Муравська та Т. Сліпченко наголошують, що «існує нагальна потреба в уніфікованому нормативно-правовому регулюванні штучного інтелекту, щоб запобігти системним і соціальним ризикам у сфері інтелектуальної безпеки» [199, с. 188]. Водночас три системні закономірності, ідентифіковані в результаті аналізу – конвергенція ринкових і ціннісних підходів, поступове ослаблення механізмів контролю під впливом лобіювання та реактивний характер регулювання – мають слугувати як орієнтирами, так і застереженнями для України. Перша закономірність набуває особливого звучання у контексті повоєнної відбудови, адже, як підкреслює М.В. Дубняк, «для повоєнної відбудови України корисно створити привабливе правове поле, що дозволить впровадити інноваційні рішення для відбудови інфраструктури та економіки України» [197, с. 41]; друга є особливо актуальним застереженням, оскільки Україна, маючи менші ресурси для незалежного технічного аудиту систем ІІІ, ще більше за ЄС ризикує потрапити у пастку саморегулювання; третя вказує на необхідність розбудови прогностичних та аналітичних спроможностей, причому А. Гачкевич справедливо наголошує на ролі

регуляторних «пісочниць», методологій оцінки ризиків, маркування та кодексів поведінки як практичних інструментів адаптації [198, с. 21]. Однак механічне перенесення норм ЄС без урахування національних особливостей – меншого обсягу ринку ШІ, обмеженості ресурсів наглядових органів та необхідності стимулювання інновацій у воєнний і післявоєнний період – вимагатиме більш гнучкого регуляторного підходу, здатного поєднати ризикоорієнтовану парадигму ЄС із реальними інституційними можливостями української держави.

Формування правової бази ЄС у сфері регулювання ШІ є прикладом поетапного та інституційно обумовленого нормотворчого процесу. Від першої Резолюції Європейського парламенту 2017 року [6] до Акту про ШІ 2024 року [25] та Плану дій «Континент ШІ» 2025 року [165] ЄС пройшов шлях від концептуалізації проблеми до створення комплексної регуляторної рамки, яка вже справляє глобальний вплив на міжнародно-правове регулювання ШІ. Подальший розвиток цієї правової бази, у тому числі через міжнародно-правові інструменти, зокрема Рамкову конвенцію Ради Європи про штучний інтелект, аналіз якої є предметом наступного підрозділу, визначатиме ефективність регулювання ШІ як на регіональному, так і на універсальному рівні.

3.2. Діяльність Ради Європи щодо захисту прав людини в умовах розвитку технологій штучного інтелекту

Проведений у попередньому підрозділі аналіз права ЄС у сфері регулювання ШІ засвідчив створення розвиненої та інституційно щільної правової бази, заснованої на ризикоорієнтованому підході та ринково-регуляторній парадигмі. Рада Європи, на відміну від ЄС, підходить до проблематики ШІ не з позицій гармонізації ринку, а з позицій захисту прав людини, демократії та верховенства права. Ця відмінність у вихідних позиціях має принципове методологічне значення: якщо Акт про ШІ є інструментом регулювання продукту (систем ШІ), то нормотворча діяльність Ради Європи спрямована на захист прав, що визначає різну

структуру актів, інструменти та механізми. Рада Європи (далі – РЄ) як міжнародна організація об'єднує 46 держав-членів і має статутну мету досягнення більшої єдності між її членами, зокрема через захист та розвиток прав людини і основоположних свобод. На відміну від ЄС, РЄ не має наднаціональних повноважень, а її нормотворча діяльність реалізується переважно через два основні інструменти: юридично обов'язкові конвенції, що потребують ратифікації державами-членами та акти «м'якого права» – рекомендації Комітету Міністрів (далі – КМ РЄ), резолюції Парламентської Асамблеї (далі – ПАРЄ), декларації та висновки спеціалізованих органів РЄ. Саме ця дуальна нормотворча архітектура визначила траєкторію діяльності РЄ у сфері ШІ: від формування консенсусу через акти «м'якого права» до розробки та прийняття юридично обов'язкового міжнародного договору, присвяченого захисту прав людини в умовах розвитку технологій ШІ.

Діяльність Ради Європи в цьому підрозділі аналізується за п'ятьма взаємопов'язаними напрямками: акти КМ РЄ та спеціалізованих органів; парламентський вимір (резолюції ПАРЄ); інституційний процес розробки Рамкової конвенції Ради Європи про штучний інтелект та права людини, демократію і верховенство права СЕТС № 225, (далі – Рамкова конвенція про ШІ); її комплексний аналіз; практика Європейського суду з прав людини у контексті технологій ШІ. Передумовою для формування актів РЄ у сфері ШІ стала вже існуюча нормативна база у сфері захисту персональних даних, ядром якої є Конвенція Ради Європи про захист осіб у зв'язку з автоматизованою обробкою персональних даних 1981 року (ETS № 108) [229], модернізована Протоколом 2018 року (СЕТС № 223), відомим як «Конвенція 108+». Модернізований Протокол запровадив принципи, безпосередньо релевантні для регулювання ШІ: прозорість (ст. 8), пропорційність (ст. 5), підзвітність та оцінку впливу (ст. 10), а також повагу до приватності за замовчуванням (ст. 10). Ключовим для проблематики ШІ є закріплене у ст. 9 Конвенції 108+ право кожної особи «не бути об'єктом рішення, що суттєво впливає на неї, яке ґрунтується виключно на автоматизованій обробці даних, без урахування

її думки», а також право на отримання знань про логіку обробки даних [229, ст. 9]. Як зазначається в аналітичному документі САНАІ «Дослідження доцільності», правова рамка, побудована навколо Конвенції 108, «залишається повністю застосовною до технології ШІ, коли оброблювані дані підпадають під сферу застосування Конвенції» [223, п. 47]. Водночас САНАІ визнав, що спеціалізований акт РЄ щодо ШІ «повинен повною мірою враховувати це *acquis*, доповнюючи його», зокрема шляхом «включення до сфери застосування таких операцій з обробки, що не стосуються виключно персональних даних, розширення сфери дії на запобігання шкоді іншим правам людини, а також включення суспільної (а не лише індивідуальної) шкоди» [223, п. 47]. Саме цю функцію покликана виконувати Рамкова конвенція про ШІ, яка, на відміну від Конвенції 108+, не обмежується лише захистом персональних даних.

Формування нормативної бази Ради Європи у сфері ШІ розпочалося з актів КМ РЄ, що створили концептуальний і термінологічний фундамент для подальшої конвенційної нормотворчості. Хронологічно першим таким актом, безпосередньо спрямованим на проблематику більш ширшого поняття алгоритмічних технологій, стала Декларація щодо маніпулятивних можливостей алгоритмічних процесів (Decl(13/02/2019)1), прийнята 13 лютого 2019 року [215]. За юридичною природою декларації КМ РЄ є актами політичного характеру, що не створюють юридичних зобов'язань, проте виражають консолідовану позицію вищого виконавчого органу РЄ та формують подальшу нормотворчу діяльність організації. Змістовна цінність цього акта полягає в ідентифікації специфічної загрози алгоритмічних процесів для когнітивної автономії індивіда: Декларація констатувала, що «сучасні інструменти машинного навчання мають зростаючу здатність не лише прогнозувати вибір, але й впливати на емоції та думки й змінювати передбачуваний курс дій, іноді підсвідомо» [215, п. 8], та визнала, що ці загрози здатні підірвати самі основи європейського правопорядку – права людини, демократію та верховенство права [215, п. 9]. Це формулювання стало концептуально новаторським, оскільки виходить за межі традиційного для РЄ дискурсу захисту персональних даних та фіксує нову категорію

загроз – маніпулятивний вплив на процес формування волі індивіда; згодом ця ідея знайшла розвиток як у резолюціях ПАРЄ, так і в ст. 5 Рамкової конвенції про ШП. Водночас Декларація має очевидні обмеження: вона не містить нормативних визначень ключових понять і обмежується закликами до держав-членів оцінити необхідність суворіших регуляторних заходів [215], не конкретизуючи ані їх характер, ані механізми імплементації. Тим не менш, як політичний сигнал Декларація виконала важливу функцію привернення уваги держав-членів до цієї проблематики і сприяла формуванню інституційного консенсусу щодо необхідності системної відповіді.

Якісно інший рівень нормативної конкретизації забезпечила Рекомендація Комітету Міністрів CM/Rec(2020)1 щодо впливу алгоритмічних систем на права людини, прийнята 8 квітня 2020 року [23]. Рекомендації КМ РЄ, прийняті на підставі ст. 15.b Статуту РЄ, є актами «м'якого права», що не мають юридично обов'язкової сили, однак адресуються безпосередньо урядам держав-членів і встановлюють стандарти належної практики, на які неодноразово посилався ЄСПЛ як на релевантне джерело для тлумачення. Рекомендація CM/Rec(2020)1 стала першим актом РЄ, що систематизував зобов'язання держав та відповідальність приватного сектору у сфері алгоритмічних систем, та запровадила робоче визначення «алгоритмічних систем» як таких, що використовують методи математичної оптимізації для збору, класифікації, виведення даних та прийняття рішень [23, Додаток, п. 2]. Це визначення відображає підхід РЄ – регулювати не конкретну технологію, а її вплив на права людини, незалежно від рівня складності алгоритму – та, як зауважує Л. Лейн, «конкретизує стандарти міжнародного права прав людини» [236, с. 935], перетворюючи абстрактні правозахисні гарантії на практичні орієнтири. Додаток до Рекомендації – Керівні принципи щодо впливу алгоритмічних систем на права людини – містить розгорнуту систему вимог: законодавче забезпечення (п. 1.1), постійний перегляд (п. 1.2), оцінку впливу на права людини (п. 5.2) та забезпечення незалежного експертного нагляду (п. 5.3). Особливо важливим є п. 15 Додатку, що закріплює принцип обережності: держави-члени повинні

«вимагати відмови від певних систем, коли їх розгортання призводить до високих ризиків незворотної шкоди або коли через їхню непрозорість людський контроль та нагляд стають непрактичними» [23, ст. 6]. Ця вимога є суттєво жорсткішою, ніж аналогічні положення Рекомендації ЮНЕСКО з етики ШІ, та фактично передбачає можливість заборони певних алгоритмічних систем – підхід, який згодом був імплементований у ст. 5 Акту про ШІ щодо заборонених практик.

Поряд із зазначеними актами, РЄ сформувала значний масив галузевих актів «м'якого права», що конкретизують правозахисні стандарти у специфічних сферах застосування ШІ. Хронологічно першою стала Європейська етична хартія щодо використання ШІ у судових системах та їхньому середовищі, прийнята Європейською комісією з ефективності правосуддя (ЄКЕП/СЕРЕЛ) 3–4 грудня 2018 року [228] – піонерський галузевий акт, що визначив п'ять основних принципів використання ШІ у правосудді (повага до основоположних прав, недискримінація, якість та безпека, прозорість та пояснюваність, контроль з боку користувача) і став прецедентом залучення спеціалізованих органів РЄ до нормотворчості у сфері ШІ. Подальші рекомендації КМ РЄ поширили правозахисний аналіз на чутливі секторальні сфери: CM/Rec(2021)8 [211] – на перетин захисту даних та автоматизованого прийняття рішень у контексті профілювання; CM/Rec(2022)11 [212] та CM/Rec(2022)13 [213] – на інформаційний простір, алгоритмічну модерацію контенту та вплив цифрових платформ на плюралізм думок; а CM/Rec(2024)5 [214], прийнята 9 жовтня 2024 року, – на особливо чутливу сферу кримінально-виконавчого права, регулюючи використання ШІ пенітенціарними та пробаційними службами.

Аналіз сукупності цих актів дозволяє ідентифікувати щонайменше три системні закономірності: послідовне розширення предметного поля – від вузької проблематики маніпулятивних можливостей алгоритмів через комплексний підхід до впливу алгоритмічних систем на права людини до галузевих стандартів у сфері правосуддя, медіа, профілювання та пенітенціарних служб; поступове підвищення нормативної конкретизації – від політичних декларацій до детальних керівних

принципів із конкретними процедурними вимогами; та формування інтертекстуальної мережі, в якій кожен наступний документ посилається на попередні та розвиває їх положення. Ця система актів виконала функцію «передконвенційного» нормативного шару, який концептуально та термінологічно підготував розробку Рамкової конвенції про ШІ і водночас продовжує зберігати самостійне значення для тлумачення та імплементації конвенційних положень.

Паралельно з діяльністю Комітету Міністрів та спеціалізованих органів, вагомий внесок у формування позиції РЄ щодо ШІ здійснила Парламентська Асамблея Ради Європи (ПАРЄ). 22 жовтня 2020 року Постійний комітет ПАРЄ, діючи від імені Асамблеї, прийняв пакет із п'яти тематичних резолюцій, кожна з яких розглядала окремий аспект впливу ШІ на права людини, демократію та верховенство права. За своєю юридичною природою резолюції ПАРЄ є актами рекомендаційного характеру, що не створюють юридичних зобов'язань для держав-членів, проте відображають консолідовану позицію національних парламентаріїв 46 держав і легітимізують подальші нормотворчі ініціативи організації.

Резолюція 2341 (2020) «Необхідність демократичного управління штучним інтелектом» [216] стала системоутворюючим актом пакету, що визначив загальну рамку парламентського бачення регулювання ШІ. Принципово важливим є те, що Резолюція визначила п'ять основних етичних принципів, які мають бути покладені в основу будь-якого регулювання ШІ: прозорість, включаючи доступність та пояснюваність; справедливість та неупередженість, включаючи недискримінацію; людська відповідальність за рішення, включаючи відповідальність та доступність засобів правового захисту; безпека та захищеність; приватність та захист даних [216, п. 14.2]. Ці п'ять принципів, сформульовані як спільний додаток до всіх п'яти резолюцій пакету, стали «етичним каркасом» підходу ПАРЄ та значною мірою вплинули на зміст Рамкової конвенції про ШІ. Резолюція також закріпила процедурну вимогу щодо створення «незалежного та проактивного механізму нагляду за участю всіх відповідних зацікавлених сторін» на базі «висококомпетентного органу (зокрема в технічному, правовому та етичному сенсі),

здатного відстежувати нові розробки у сфері цифрових технологій та точно й авторитетно оцінювати їхні ризики та наслідки» [216, п. 15] – вимога, що лише частково була реалізована у фінальному тексті Рамкової конвенції про ШІ, яка обмежилася загальним положенням ст. 26 про необхідність створення механізмів нагляду без конкретизації їхнього інституційного формату.

Резолюція 2342 (2020) «Правосуддя за алгоритмом – Роль штучного інтелекту у поліцейській діяльності та системах кримінального правосуддя» [217] зосередилася на одній із найбільш чутливих сфер застосування ШІ. Резолюція ідентифікувала конкретні застосування ШІ, що вже використовувалися або планувалися до впровадження в державах-членах РЄ (розпізнавання обличчя, предиктивна поліцейська діяльність, оцінка ризиків при прийнятті рішень щодо тримання під вартою та призначення покарання), та наголосила, що використання ШІ у поліцейській діяльності та кримінальному правосудді «може мати значні переваги, якщо воно належним чином регулюється, але ризикує мати особливо серйозний вплив на права людини, якщо це не так» [217, п. 8]. Серед конкретних вимог Резолюції варто виокремити заклик до держав-членів «підтримувати реєстр усіх застосувань ШІ, що використовуються у публічному секторі, та звертатися до нього при розгляді нових застосувань, щоб ідентифікувати та оцінити можливі кумулятивні впливи» [217, п. 9.2]. Ця вимога фіксує проблему кумулятивного ефекту – ситуації, коли «застосування систем ШІ в окремих, але пов'язаних контекстах, особливо різними відомствами, які послідовно покладаються на роботу одне одного, може мати несподівані, навіть непередбачувані кумулятивні впливи» [217, п. 7.5]. Ця проблема залишається недостатньо відображеною як у Рамковій конвенції про ШІ, так і в Акті про ШІ, що свідчить про випередження парламентського дискурсу відносно нормотворчої практики.

Резолюція 2343 (2020) «Запобігання дискримінації, спричиненій використанням штучного інтелекту» [218] сфокусувалася на проблемі алгоритмічної упередженості та її впливі на права рівності та недискримінації. Резолюція констатувала, що «алгоритми, оптимізовані для ефективності, прибутковості або

інших цілей, без належного врахування необхідності гарантувати рівність та недискримінацію, можуть спричиняти пряму або непряму дискримінацію, включаючи дискримінацію за асоціацією, за широким спектром ознак» [218, п. 4], та звернула особливу увагу на проблему доказування дискримінації в контексті алгоритмічного прийняття рішень: «складність систем ШІ та той факт, що вони часто розробляються приватними компаніями і розглядаються як їхня інтелектуальна власність», робить «дискримінацію надзвичайно складною для доказування» [218, п. 5].

Резолюції 2344 (2020) «Інтерфейс мозок-комп'ютер: нові права чи нові загрози для основоположних свобод?» [219] та 2345 (2020) «Штучний інтелект та ринки праці: друг чи ворог?» [220] доповнили тематичний спектр, охопивши проблематику нейротехнологій та впливу ШІ на зайнятість, хоча їх практичний вплив на подальшу нормотворчість виявився менш значущим. Прийнята у 2023 році Резолюція 2485 (2023) «Поява летальних автономних систем зброї (LAWS) та необхідність їхнього осмислення через призму європейського права прав людини» [221] поширила правозахисний аналіз на воєнну сферу, закликала до заборони повністю автономних систем зброї та наполягала на збереженні «значущого людського контролю» над застосуванням сили – позиція, яка кореспондує з принципом людського нагляду, закріпленим у ст. 8 Рамкової конвенції про ШІ.

Кульмінацією парламентського виміру діяльності РЄ у сфері ШІ став Висновок 303 (2024) ПАРЄ щодо проєкту Рамкової конвенції про ШІ, прийнятий 18 квітня 2024 року [222]. За юридичною природою висновок ПАРЄ є обов'язковою процедурною стадією прийняття конвенцій РЄ відповідно до ст. 22 Статуту організації, що надає йому критичне значення для забезпечення демократичної легітимації конвенції. У Висновку ПАРЄ підтвердила підтримку конвенції, нагадавши про «набір ключових етичних принципів» зі своїх попередніх резолюцій та «тверду впевненість у тому, що правове регулювання є необхідним для уникнення або пом'якшення потенційних ризиків для демократії, прав людини та верховенства

права, що виникають у зв'язку з використанням ШІ» [222, п. 3], – і завершив парламентський цикл, розпочатий пакетом резолюцій 2020 року.

Оцінюючи парламентський вимір діяльності РЄ загалом, слід зазначити, що резолюції ПАРЕ виконали кілька важливих функцій у процесі формування підходів РЄ до регулювання ШІ: сформулювали набір етичних принципів, що став консенсусною основою для подальших переговорів; ідентифікували конкретні проблемні сфери – від кримінального правосуддя до ринків праці; створили політичний мандат для розробки юридично обов'язкового інструменту; та виявилися амбітнішими за фінальний текст Рамкової конвенції про ШІ в окремих аспектах – зокрема щодо вимог до інституційного нагляду, кумулятивних ефектів алгоритмічних систем та реєстру застосувань ШІ у публічному секторі, що свідчить про певне зниження рівня правозахисних гарантій у ході міжурядових переговорів. Масштаб та конкретні прояви цього зниження, простежені нижче на прикладі порівняння нульового проєкту з фінальним текстом конвенції, засвідчують структурну закономірність, аналогічну тій, що спостерігалася при розробці Акту про ШІ і була проаналізована раніше.

Паралельно з парламентським дискурсом, РЄ ініціювала масштабний інституційний процес розробки юридично обов'язкового договору – процес, що тривав понад п'ять років і завершився прийняттям Рамкової конвенції про ШІ. 11 вересня 2019 року Комітет Міністрів створив Спеціальний міжурядовий комітет зі штучного інтелекту (САНАІ) та уповноважив його дослідити доцільність та потенційні елементи правової рамки щодо ШІ, заснованої на стандартах РЄ у сфері прав людини, демократії та верховенства права [223]. Мандат САНАІ відображав обережний підхід: комітету було доручено не розробку конвенції, а лише дослідження доцільності такої розробки, що свідчить про відсутність на той момент консенсусу серед держав-членів щодо необхідності юридично обов'язкового інструменту. Головним результатом діяльності САНАІ став аналітичний звіт «Дослідження доцільності» (САНАІ(2020)23), завершений 17 грудня 2020 року [223], який проаналізував існуючі правові рамки, ідентифікував прогалини та

обґрунтував необхідність юридично обов'язкового інструменту, наголосивши, що існуючі міжнародно-правові акти, хоча й застосовні до контексту ШІ, не були спеціально розроблені для вирішення унікальних правових викликів систем ШІ, а Рада Європи має «унікальні можливості для того, щоб очолити ці зусилля, та – спираючись на існуючі рамки – спрямувати узгодження систем ШІ з її стандартами щодо прав людини, демократії та верховенства права» [223].

На основі «Дослідження доцільності» САНАІ підготував другий ключовий документ – «Можливі елементи правової рамки щодо штучного інтелекту, заснованої на стандартах Ради Європи щодо прав людини, демократії та верховенства права» (САНАІ(2021)08rev), прийнятий 9 грудня 2021 року [224], який окреслив концептуальну архітектуру майбутнього акту – основні тематичні блоки та принципи юридично обов'язкового договору. У 2022 році Комітет Міністрів трансформував САНАІ у Комітет зі штучного інтелекту (CAI), надавши йому мандат на створення правового інструменту, що «базується на стандартах Ради Європи щодо прав людини, демократії та верховенства права, і сприяє інноваціям» [227, с. 1]. Зміна мандату – від дослідження доцільності до безпосередньої розробки конвенції – засвідчила досягнення консенсусу щодо необхідності юридично обов'язкового інструменту, а сама формула «сприяє інноваціям» відображала компромісну позицію між правозахисними та промисловими інтересами – компроміс, аналогічний тому, що спостерігався при розробці Акту про ШІ.

Процес розробки тексту Конвенції пройшов кілька послідовних етапів. 6 січня 2023 року був представлений Переглянутий нульовий проєкт (CAI(2023)01) [225]. Цей документ окреслив базову структуру Конвенції та запропонував формулювання ключових положень. 7 липня 2023 року, після першого читання нульового проєкту, був представлений Консолідований робочий проєкт (CAI(2023)18) [226], що містив положення, попередньо узгоджені під час першого читання, а також пропозиції Голови щодо неузгоджених питань. Порівняння нульового та фінального текстів Конвенції свідчить про суттєву еволюцію тексту, що має характер систематичного звуження обсягу зобов'язань та розширення дискреції держав-учасниць. Цю

тенденцію можна простежити щонайменше за чотирма ключовими параметрами. По-перше, і найсуттєвіше, ст. 4(2) нульового проєкту встановлювала, що Конвенція застосовується до систем ШІ «незалежно від того, чи здійснюється ця діяльність публічними або приватними суб'єктами» [225, ст. 4(2)], а ст. 8, що мала назву «Зобов'язання стосовно публічних та приватних суб'єктів», безпосередньо покладала на обидві категорії суб'єктів обов'язок забезпечити сумісність застосування ШІ з правами людини. Фінальний текст замінив цю пряму та універсальну застосовність дуальною моделлю зі значною дискрецією держав щодо приватного сектору (ст. 3(1)(b)). По-друге, нульовий проєкт містив окрему ст. 11 «Збереження громадського здоров'я та довкілля», що зобов'язувала кожну Сторону «вжити необхідних заходів для збереження громадського здоров'я та довкілля у контексті застосування системи штучного інтелекту» [225, ст. 11], – цей екологічний вимір був повністю вилучений із фінального тексту. По-третє, ст. 14 нульового проєкту закріплювала «Принцип підзвітності, відповідальності та юридичної відповідальності» [225, ст. 14], тоді як ст. 9 фінального тексту обмежилася підзвітністю та відповідальністю, опустивши пряме посилення на юридичну відповідальність – елемент, принципово важливий для забезпечення ефективного захисту прав осіб, що постраждали від систем ШІ. По-четверте, нульовий проєкт містив вужче виключення – лише для «національної оборони» (ст. 4(3)), тоді як фінальний текст розширив його до «національної безпеки» (ст. 3(2)), що є значно ширшою категорією, здатною охопити спостереження, розвідувальну діяльність та інші чутливі сфери застосування ШІ.

Суттєвою особливістю переговорного процесу, що значною мірою пояснює описану траєкторію, стала участь у розробці конвенції позаєвропейських держав. Як зазначають М.А. Пресно Лінера та А. Мьовезе, «позаєвропейські держави були включені в переговори у процесі розробки Рамкової конвенції (Аргентина, Австралія, Канада, Коста-Ріка, Святий Престол, Ізраїль, Японія, Мексика, Перу, Сполучені Штати Америки та Уругвай)» [232, с. 309], а також представники інших міжнародних організацій – ОБСЄ, ОЕСР та ЮНЕСКО. Ця інклюзивність

переговорного процесу відображає глобальні амбіції Конвенції та відрізняє її від Акту про ШІ, який розроблявся виключно в рамках законодавчих процедур ЄС. Водночас участь таких держав, як США, які традиційно скептично ставляться до обов'язкового регулювання приватного сектору у сфері ШІ, неминуче вплинула на зниження рівня зобов'язань порівняно з нульовим проєктом. Ця закономірність є структурно аналогічною тій, що була ідентифікована раніше при аналізі Акту про ШІ, в якому механізми зовнішнього контролю поступово замінювалися самооцінкою та саморегулюванням. Різниця полягає у джерелі тиску: якщо в ЄС ключовим чинником ослаблення стала позиція технологічної індустрії, то у випадку Рамкової конвенції визначальну роль відіграла позиція позаєвропейських держав-учасниць переговорного процесу.

Рамкова конвенція Ради Європи про штучний інтелект та права людини, демократію і верховенство права CETS № 225 (Рамкова конвенція про ШІ) є другим після Акту про ШІ юридично обов'язковим інструментом регулювання ШІ, прийнятим на регіональному рівні. Як зазначають М.А. Пресно Лінера та А. Мьовезе, «Акт про ШІ та Рамкова конвенція можуть розглядатися як європейські відповіді на величезний виклик просування інновацій та конкурентоспроможності у сфері ШІ способом, сумісним із захистом основоположних прав та верховенством права» [232, с. 293] – проте з принципово різною юридичною технікою: якщо ЄС обрав детальне горизонтальне регулювання з конкретними зобов'язаннями операторів та системою санкцій, то РЄ – встановлення правозахисних стандартів з делегуванням конкретної імплементації державам-учасницям. Конвенція була прийнята КМ РЄ 17 травня 2024 року та відкрита для підписання 5 вересня 2024 року у Вільнюсі [24] як міжнародний договір, обов'язковий для держав-учасниць та відкритий для приєднання не лише членам РЄ та ЄС, а й іншим державам. Її рамковий характер, прямо зафіксований у назві та преамбулі, означає встановлення загальних принципів із значною свободою імплементації для держав та з можливістю доповнення «подальшими інструментами для вирішення конкретних питань, пов'язаних з діяльністю протягом життєвого циклу систем штучного

інтелекту» [24]. Ст. 1 визначає об'єкт Конвенції – забезпечення того, щоб «діяльність протягом життєвого циклу систем штучного інтелекту була повністю сумісною з правами людини, демократією та верховенством права» [24, ст. 1(1)], а ст. 2 запроваджує визначення системи ШІ як «машинної системи, яка для явних або неявних цілей виводить з отриманих вхідних даних спосіб генерування таких вихідних даних, як прогнози, контент, рекомендації або рішення, що можуть впливати на фізичне або віртуальне середовище» [24, ст. 2]. Це визначення, як зазначають М.А. Пресно Лінера та А. Мьовезе, стало результатом «зусиль в останню мить з метою узгодження Акту про ШІ та Рамкової конвенції настільки, наскільки це можливо» [232, с. 309] і ґрунтується на функціональному підході, заснованому на понятті «виведення» (англ. – *inference*), закріпленому у переглядах Рекомендації ОЕСР листопада 2023 та травня 2024 року [9]. Водночас між цими регіональними актами зберігається й суттєва відмінність: визначення Акту про ШІ додатково включає критерії «функціонування з різним ступенем автономії» та «адаптивності після розгортання» [25, ст. 3(1)], тоді як Рамкова конвенція свідомо уникає цих критеріїв, зберігаючи більш загальне формулювання, що відповідає її рамковому характеру та необхідності прийнятності визначення для держав-учасниць з різними правовими традиціями.

Особливої уваги заслуговує ст. 3 Конвенції, що визначає сферу її застосування та закріплює дуальну модель, еволюція якої від універсальної застосовності нульового проєкту до диференційованого підходу фінального тексту була детально простежена вище. Щодо публічного сектору Конвенція встановлює прямі та безумовні зобов'язання: «кожна Сторона застосовує цю Конвенцію до діяльності протягом життєвого циклу систем штучного інтелекту, що здійснюється публічними органами влади або приватними суб'єктами, що діють від їхнього імені» [24, ст. 3(1) (a)]. Щодо приватного сектору кожна Сторона «вирішує ризики та наслідки, що виникають у зв'язку з діяльністю протягом життєвого циклу систем штучного інтелекту приватних суб'єктів... у спосіб, що відповідає об'єкту та меті цієї Конвенції» [24, ст. 3(1)(b)], із поданням декларації щодо способу імплементації. Ця

конструкція, що є результатом компромісу з позаєвропейськими учасниками переговорного процесу, створює ризик фрагментації та асиметрії захисту, оскільки рівень гарантій для осіб, чії права порушуються приватними суб'єктами, залежатиме від обраного кожною державою способу імплементації. Цю критику чітко висловила ПАРЄ у Висновку 303 (2024), констатувавши, що такий підхід «далекий від ідеального з точки зору правової визначеності та передбачуваності зобов'язань», «не відповідає позиціям, раніше висловленим Асамблеєю, Комісаром Ради Європи з прав людини та САНАІ», суперечить принципу позитивних зобов'язань держав щодо захисту осіб від порушень прав людини з боку приватних суб'єктів та «створює суттєву прогалину» [222, § 7]. Ця оцінка є принципово значущою, оскільки переважна більшість систем ШІ, що безпосередньо впливають на права та свободи індивідів, розробляються та розгортаються саме приватними суб'єктами. Аналогічні застереження стосуються й виключення для національної безпеки (ст. 3(2)), обсяг якого суттєво розширився порівняно з нульовим проектом: попри застереження про сумісність із «застосовним міжнародним правом, включаючи міжнародне право прав людини» [24], саме у сфері національної безпеки (спостереження, розвідувальна діяльність, автономні системи зброї) ризики для прав людини є найбільш гострими, а абстрактне посилення на міжнародне право без конкретизації процедурних гарантій та механізмів контролю може виявитися недостатнім для запобігання зловживанням.

Матеріально-правовий зміст Конвенції структуровано за кількома тематичними блоками. Глава II (ст. 4–5) встановлює загальні зобов'язання щодо захисту прав людини та демократичних інститутів; зокрема, ст. 5 зобов'язує кожную Сторону «ухвалювати або підтримувати заходи, спрямовані на забезпечення того, щоб системи ШІ не використовувалися для підриву цілісності, незалежності та ефективності демократичних інститутів і процесів, включаючи принцип поділу влади, повагу до незалежності судової влади та доступ до правосуддя» [24, ст. 5(1)]. Глава III (ст. 6–13) формулює загальні принципи, що їх кожна Сторона має імплементувати «у спосіб, що відповідає її внутрішній правовій системі» [24, ст. 6]:

людська гідність та індивідуальна автономія (ст. 7), прозорість та нагляд (ст. 8), підзвітність та відповідальність (ст. 9), рівність та недискримінація (ст. 10), приватність та захист персональних даних (ст. 11), надійність (ст. 12) і безпечні інновації (ст. 13). Ці принципи значною мірою відтворюють етичний каркас, сформульований у п'яти резолюціях ПАРЄ 2020 року та Рекомендаціях КМ, проаналізованих вище, і за рамковим характером Конвенції не конкретизуються до рівня деталізованих зобов'язань.

Більш детального аналізу потребує Глава IV «Засоби правового захисту» (ст. 14–15), яка, на відміну від принципів Глави III, покликана забезпечити практичну можливість захисту прав осіб, постраждалих від систем ШІ, і тому є визначальною для оцінки ефективності всього конвенційного механізму. Ст. 14(1) зобов'язує кожну Сторону «в обсязі, в якому засоби правового захисту вимагаються її міжнародними зобов'язаннями та відповідно до її внутрішньої правової системи, ухвалювати або підтримувати заходи для забезпечення доступності ефективних засобів правового захисту у разі порушень прав людини, що випливають з діяльності протягом життєвого циклу систем штучного інтелекту» [24], передбачаючи зокрема документування систем ШІ та надання інформації постраждалим особам, її достатності для оскарження відповідних рішень, а також «ефективну можливість для зацікавлених осіб подати скаргу до компетентних органів» [24, ст. 14(2)(с)]. Ст. 15 доповнює цей механізм вимогою забезпечити «ефективні процедурні гарантії, захисні заходи та права» для осіб, на реалізацію прав яких система ШІ «суттєво впливає» [24, ст. 15(1)]. Водночас аналіз цих положень виявляє кілька системних проблем. По-перше, застереження «в обсязі, в якому засоби правового захисту вимагаються її міжнародними зобов'язаннями» фактично не створює автономного стандарту, а відсилає до вже існуючих вимог, зокрема ст. 13 Конвенції про захист прав людини і основоположних свобод 1950 року (Європейська конвенція з прав людини, далі – ЄКПЛ), яка, як буде показано нижче в аналізі практики ЄСПЛ, сама потребує адаптації до контексту алгоритмічного спостереження, систем ШІ та автоматизованого прийняття рішень. По-друге, ст. 14(2)(с) обмежується

формулюванням про «компетентні органи», не конкретизуючи їх судового характеру, – саме цю прогалину зафіксувала ПАРЄ у Висновку 303 (2024), запропонувавши додати посилання на «судові органи» або «судовий перегляд» [222, п. 9.5]. По-третє, Конвенція не містить положень щодо юридичної відповідальності за шкоду, завдану системами ШІ, хоча ст. 14 Нульового проєкту прямо передбачала «юридичну відповідальність» (англ. – *legal liability*) поряд із підзвітністю та відповідальністю, що створює розрив між визнанням права на засіб правового захисту та відсутністю матеріально-правової основи для його реалізації.

Порівняння структури зобов'язань Рамкової конвенції та Акту про ШІ виявляє принципові відмінності у регуляторному підході. Якщо Акт про ШІ побудований на детальній класифікації систем ШІ за рівнями ризику із конкретними вимогами для кожної категорії, Конвенція оперує загальними принципами, залишаючи конкретизацію на розсуд держав-учасниць. Як зазначають М.А. Пресно Лінера та А. Мьовезе, «Рамкова конвенція лише передбачає, що кожна Сторона створює або визначає один або кілька ефективних механізмів для нагляду за дотриманням зобов'язань за цією Конвенцією» [232, с. 307], тоді як Акт про ШІ містить розгорнуту систему санкцій, включаючи значні адміністративні штрафи. Ця структурна відмінність відображає різну юридичну природу двох інструментів: регламент ЄС є актом прямої дії, тоді як рамкова конвенція потребує імплементації через національне законодавство. Водночас обидва інструменти є взаємодоповнюваними. Рішення Ради ЄС (EU) 2024/2218 від 28 серпня 2024 року про підписання Конвенції від імені Європейського Союзу прямо передбачає, що «Конвенція має бути імplementована в Союзі виключно через Регламент (ЄС) 2024/1689 та інший відповідний *acquis* Союзу, де це застосовно» [232, с. 310]. Це положення означає, що для держав-членів ЄС Конвенція не створює додаткових зобов'язань понад ті, що встановлені Актом про ШІ. Для держав, які не є членами ЄС, Рамкова конвенція про ШІ натомість виконує іншу функцію – як зазначають М.А. Пресно Лінера та А. Мьовезе, вона «може відігравати важливу роль як орієнтир для внутрішніх регуляторних та правових рамок» [232, с. 310]. Це положення є особливо значущим

для України, яка, не будучи членом ЄС, може використовувати Конвенцію як правову рамку для формування національного законодавства щодо ШІ, водночас орієнтуючись на більш деталізовані вимоги Акту про ШІ у контексті виконання зобов'язань за Угодою про асоціацію.

Механізм імплементації Конвенції заснований на Конференції Сторін (ст. 23), яка наділяється повноваженнями щодо моніторингу виконання, сприяння обміну інформацією та найкращими практиками, формулювання рекомендацій щодо імплементації та розгляду пропозицій про внесення змін. Ст. 24 встановлює звітний механізм: кожна Сторона має надати звіт Конференції Сторін протягом перших двох років після набуття статусу учасниці та періодично надалі. Цей механізм є значно м'якшим порівняно як із системою індивідуальних скарг ЄСПЛ, так і з системою санкцій Акту про ШІ. Л. МакГрегор та Д. Маррей справедливо зазначають, що «міжнародне право прав людини вже надає рамку» для алгоритмічної підзвітності, яка «вимагає від держав створити систему, що запобігає порушенням прав людини, встановлює механізми моніторингу та нагляду як гарантії, притягує відповідальних до відповідальності та забезпечує засіб правового захисту для осіб та груп, які стверджують, що їхні права були порушені» [233, с. 311]. Рамкова конвенція про ШІ лише частково реалізує цей потенціал, що залишає значний простір для її подальшого розвитку та конкретизації.

Невід'ємною складовою правозахисної діяльності Ради Європи є практика Європейського суду з прав людини (ЄСПЛ). Хоча Суд ще не розглядав справу, в якій безпосередньо оцінювалася б система ШІ як така, низка його рішень створила прецедентну базу, що поступово адаптується до цифрових реалій. Як зазначає М. Сапаньош, «ЄСПЛ досі не мав справи, яка була б тісно пов'язана зі ШІ», і лише в одному рішенні з приблизно 25 000 штучний інтелект взагалі згадується [235, с. 105, 107], проте «для ЄСПЛ не видається можливим довго уникати справ, пов'язаних зі ШІ» [235, с. 106]. Основним конвенційним положенням, через яке ЄСПЛ розглядає проблематику цифрового спостереження та обробки даних стала ст. 8 ЄКПЛ. Рішення Великої палати ЄСПЛ у справі *S. and Marper v. the United Kingdom* (2008)

[203] встановило фундаментальний принцип непропорційності безстрокового зберігання біометричних даних осіб, яких було виправдано, підкресливши «ризик стигматизації, що випливає з того факту, що особи у становищі заявників, які не були засуджені за жодне правопорушення та мають право на презумпцію невинуватості, отримують таке саме поводження, як і засуджені особи» [203, § 122]. Цю лінію практики продовжило рішення у справі *Tonchev v. Bulgaria* (2024) [275], в якому Суд констатував порушення ст. 8 ЄКПЛ у зв'язку з тривалим зберіганням персональних даних заявника, підкресливши, що відповідне болгарське законодавство було недостатньо передбачуваним, а розбіжності у його тлумаченні поширювалися й на застосування права ЄС у сфері захисту даних. У справі *Szabó and Vissy v. Hungary* (2016) [210] Суд визнав, що «гарантії, передбачені існуючою прецедентною практикою Конвенції щодо перехоплення комунікацій, потребують посилення для вирішення питання» масового спостереження [210, § 70], фактично констатувавши недостатність традиційних гарантій в умовах технологічного прогресу. Етапне рішення Великої палати у справі *Big Brother Watch and Others v. the United Kingdom* (2021) [204] розвинуло цей підхід, запровадивши тест «глобальної оцінки» [204, § 360] режимів масового перехоплення з фокусом на «наскрізних гарантіях» (англ. – *end-to-end safeguards*) [204, § 350], що враховують зростаючий ступінь втручання на різних стадіях обробки даних. Це рішення є безпосередньо релевантним для проблематики ШІ, оскільки масове перехоплення комунікацій дедалі більше покладається на алгоритмічні системи. Як підкреслює Е. Коста, «використання ШІ у спостереженні відкриває двері до нової ери державного спостереження, а саме алгоритмічного спостереження» [234, с. 212].

Рішення у справі *Glukhin v. Russia* (2023) [207] стало першим, у якому ЄСПЛ безпосередньо розглянув використання технології розпізнавання обличчя – для ідентифікації та арешту учасника мирного протесту. Суд констатував порушення статей 8 і 10 ЄКПЛ, визнавши «стримуючий ефект» цієї технології на свободу вираження поглядів та зібрань [207, § 88], а також наголосивши на необхідності «детальних правил, що регулюють обсяг та застосування» технології розпізнавання

обличчя та «ще більшій потребі у гарантіях, коли використовується технологія розпізнавання обличчя в режимі реального часу» [207, § 82]. Це рішення продовжило лінію, започатковану у справах *Gaughran v. the United Kingdom* (2020) [208] (визнання зберігання фотографій у поєднанні з можливістю застосування технологій розпізнавання обличчя втручанням у право на приватне життя) та *Peck v. the United Kingdom* (2003) [209] (порушення права на приватність публічним розкриттям записів відеоспостереження навіть у публічних місцях).

Рішення Великої палати у справі *Bărbulescu v. Romania* (2017) [205] розширило правозахисний дискурс на горизонтальний вимір – відносини між приватним роботодавцем та працівником. Велика палата сформулювала шість критеріїв оцінки правомірності моніторингу комунікацій працівників: попереднє повідомлення, обсяг та ступінь втручання, легітимність підстав, наявність менш інтрузивних засобів, наслідки для працівника та належність гарантій [205, § 121]. Ці критерії набувають нового виміру в умовах поширення алгоритмічних систем управління персоналом, а сама справа є релевантною для тлумачення Рамкової конвенції про ШІ, оскільки ілюструє проблему у приватному секторі – сфері, щодо якої Конвенція надає державам-учасницям найбільшу дискрецію.

Аналіз практики ЄСПЛ дозволяє констатувати, що Суд послідовно тлумачить конвенційні права у світлі технологічного розвитку відповідно до доктрини «живого інструменту», однак, як наголошує І.М. Берназюк, «практика ЄСПЛ має переважно реактивний характер – Суд реагує на конкретні скарги, що надходять до нього, а це означає певне відставання від темпів технологічного розвитку» [238, с. 95]. Цей висновок про реактивний характер практики ЄСПЛ поділяють Н. Шишкова та М. Медведєва, які, проаналізувавши низку справ про захист даних і цифрові технології, зазначають, що «ЄСПЛ не мав нагоди безпосередньо проаналізувати вплив ШІ на права людини у своїй практиці» [290, с. 53]. Навіть у справі *Glukhin v. Russia*, що стосувалася технології розпізнавання облич, Суд, як зазначають авторки, не розглядав ШІ як окрему правову категорію. Цю тезу підтверджує суддя ЄСПЛ А. Бордсен, констатуючи, що «жодне з рішень Суду на сьогодні, яке можна було б

вважати ключовим щодо цифрового/технологічного прогресу, не стосується штучного інтелекту як такого» [237], водночас окреслюючи загрози, пов'язані з ШІ: ерозію права на приватність, структурну шкоду людській гідності, поглиблення дискримінації, виклики свободі вираження через алгоритмічну модерацію контенту та загрози цілісності виборчого процесу. Показовим у цьому контексті є те, що Судовий семінар ЄСПЛ, проведений 31 січня 2025 року з нагоди відкриття судового року, був присвячений темі «Захист прав людини у світі штучного інтелекту, алгоритмів та великих даних», що засвідчує цілеспрямовану підготовку Суду до розгляду майбутніх справ зі ШІ. Як зазначають М.А. Пресно Лінера та А. Мьовезе, Рамкова конвенція про ШІ, ймовірно, «функціонуватиме як інтерпретативний документ для Європейського суду з прав людини» [232, с. 307], що дозволить наповнити конкретним змістом абстрактні конвенційні гарантії ЄКПЛ у контексті ШІ.

Узагальнюючи проведений аналіз діяльності РЄ у сфері захисту прав людини в умовах розвитку технологій ШІ, можемо констатувати формування цілісної правової бази, що охоплює акти «м'якого права» (резолюції ПАРЕ, декларації та рекомендації КМ РЄ), «твердого права» (Рамкова конвенція про ШІ) та судову практику ЄСПЛ. Специфіка підходу РЄ – правозахисна парадигма – принципово відрізняє його від ринково-регуляторної парадигми ЄС та парадигми сталого розвитку ООН: якщо ООН концентрується на глобальній координації та подоланні цифрового розриву, а ЄС – на гармонізації внутрішнього ринку та конкурентоспроможності, то РЄ фокусується на забезпеченні того, щоб використання ШІ не підривало фундаментальні цінності європейського правового порядку. Водночас проведений аналіз виявляє чотири суттєві виклики: розрив між амбіційністю принципів та обмеженістю механізмів їх забезпечення, оскільки звітність перед Конференцією Сторін є значно м'якшим механізмом порівняно як із системою індивідуальних скарг ЄСПЛ, так і з системою санкцій Акту про ШІ; дуальну модель сфери застосування, що надає державам-учасницям надмірну дискрецію щодо регулювання приватного сектору та ризикує створити асиметрію

захисту; виключення для національної безпеки, яке потенційно виводить з-під конвенційного контролю найбільш чутливі застосування ІІІ; та реактивний характер судової практики ЄСПЛ, яка неминуче відстає від технологічного розвитку.

Для України значення діяльності РЄ у сфері регулювання ІІІ є багатовимірним. 15 травня 2025 року Україна підписала Рамкову конвенцію про ІІІ [230], супроводивши підписання двома деклараціями [231]: перша, подана відповідно до ст. 32(1) Конвенції, передбачає незастосування Конвенції до тимчасово окупованих рф територій України до повного припинення збройної агресії та відновлення територіальної цілісності, а друга, подана відповідно до ст. 3(1)(b) Конвенції, фіксує обрання Україною моделі «інших належних заходів» щодо діяльності приватних суб'єктів, що надає регуляторну гнучкість, проте створює зобов'язання щодо розробки національного законодавства, яке має забезпечити рівень захисту прав людини, еквівалентний конвенційним стандартам. Подальша ратифікація конвенції залишається стратегічним пріоритетом, оскільки лише вона перетворить політичний акт підписання на юридично обов'язкове зобов'язання. При цьому конвенція набуває для України значення передусім як концептуальна рамка для формування національного законодавства про ІІІ, відсутність якого є очевидною прогалиною: Концепція розвитку штучного інтелекту в Україні, схвалена розпорядженням Кабінету Міністрів від 2 грудня 2020 р. № 1556-р [49], визначає загальні стратегічні напрями, однак досі не отримала подальшого втілення в окремих галузях національного права.

Окремі положення конвенції набувають особливої актуальності з огляду на специфічні виклики, з якими стикається Україна. Ст. 5, що зобов'язує Сторони «ухвалювати або підтримувати заходи, спрямовані на забезпечення того, щоб системи штучного інтелекту не використовувалися для підриву цілісності, незалежності та ефективності демократичних інститутів і процесів» [24, ст. 5(1)], є безпосередньо релевантною в контексті систематичного використання ІІІ-генерованої дезінформації як інструменту гібридної війни – «Дослідження

доцільності» САНАІ фіксувало, що «інструменти на основі ШІ посилили масштаб та охоплення» пропаганди, «уможливлюючи масштабні, але водночас цілеспрямовані дезінформаційні кампанії через координовану неавтентичну поведінку, зокрема через дідфейковий контент» [223, п. 35]. Особливе значення для українського контексту має також: ст. 10 щодо рівності та недискримінації (в умовах цифровізації державних послуг через платформу «Дія» та алгоритмічного розподілу соціальних виплат); ст. 8 щодо прозорості, включно з «ідентифікацією контенту, створеного системами штучного інтелекту» [24, ст. 8] (у контексті використання технологій розпізнавання облич під час збройного конфлікту); ст. 14 щодо ефективних засобів правового захисту (у кореляції з гарантіями ст. 9 Конвенції 108+ [229] стосовно інформації про логіку алгоритмічного прийняття рішень); та ст. 16 щодо оцінки та управління ризиками (в умовах стрімкого впровадження ШІ у воєнній, правоохоронній сфері та державному управлінні без належної попередньої оцінки впливу на права людини). Рамкова конвенція про ШІ може слугувати інтегративною правовою основою для формування національного законодавства щодо ШІ, забезпечуючи одночасну відповідність як правозахисним стандартам РЄ, так і вимогам *acquis* ЄС у контексті виконання Угоди про асоціацію.

Подальший розвиток правового регулювання ШІ в межах РЄ визначатиметься кількома факторами – темпами ратифікації Рамкової конвенції про ШІ, ефективністю Конференції Сторін як моніторингового механізму, еволюцією практики ЄСПЛ у контексті нових технологій та здатністю системи «м'якого права» адаптуватися до стрімкого технологічного розвитку. Порівняльний аналіз регуляторних підходів інших регіональних організацій, який є предметом наступного підрозділу, дозволить оцінити місце актів ЄС та РЄ у ширшому контексті міжнародно-правового регулювання ШІ та виявити спільні і відмінні риси в ньому.

3.3. Інші регіональні ініціативи у сфері міжнародно-правового регулювання штучного інтелекту: порівняльний аналіз

Проведений у попередніх підрозділах аналіз правової бази ЄС та Ради Європи у сфері регулювання ШІ засвідчив формування розвиненої регуляторної моделі: ЄС створив перший у світі юридично обов'язковий акт для держав-членів – Регламент (ЄС) 2024/1689 (Акт про ШІ), тоді як Рада Європи прийняла першу юридично обов'язкову міжнародну конвенцію – Рамкову конвенцію про штучний інтелект та права людини, демократію і верховенство права. Водночас ці два міжнародно-правові акти, попри їхню безперечну нормативну значущість, не вичерпують усього спектру правового регулювання ШІ, що формується на різних континентах та відображає істотно відмінні ціннісні, економічні та інституційні пріоритети. Поліцентричний характер міжнародно-правового регулювання ШІ, виявлений у Розділі 2, передбачає існування множинних центрів нормотворчості, кожен з яких розробляє власні підходи; переважна більшість населення планети та держав знаходяться поза межами регуляторної системи ЄС; а порівняльний аналіз різних регіональних моделей дозволяє виявити як спільні закономірності, так і специфічні особливості кожного регіону.

Предметом аналізу в цьому підрозділі є ініціативи з міжнародно-правового регулювання ШІ на рівні чотирьох регіонів за межами Європи: Азійсько-Тихоокеанського, Африканського, Американського та Арабського. На відміну від стандартної класифікації ООН, що виділяє п'ять регіонів (Африка, Америка, Азія, Європа, Океанія), наша типологія побудована не суто за географічним, а за інституційно-функціональним критерієм – наявністю активних регіональних міжнародних організацій, що прийняли формалізовані нормативні акти у сфері регулювання ШІ. Океанія не виділяється окремо, оскільки регіональні структури станом на 2026 рік не ухвалили спеціалізованих актів щодо ШІ. Натомість Арабський регіон, який за класифікацією ООН розподілений між Африкою та Азією, виділяється як самостійна одиниця аналізу завдяки наявності спільних міжнародних організацій, що прийняли власні акти у сфері регулювання ШІ. Національне законодавство окремих держав не є предметом дослідження, оскільки належить до сфери внутрішнього публічного права, хоча окремі національні ініціативи можуть

згадуватися як ілюстрація взаємодії регіонального та національного рівнів. Хронологічні рамки аналізу охоплюють період 2019–2026 років.

Азійсько-Тихоокеанський регіон є одним із найдинамічніших центрів формування правового регулювання ШІ за межами Європи та характеризується інституційною поліцентричністю: питання ШІ одночасно перебувають у полі зору Асоціації держав Південно-Східної Азії (АСЕАН), Азійсько-Тихоокеанського економічного співробітництва (АТЕС) та Шанхайської організації співробітництва (ШОС). Кожна з цих організацій репрезентує відмінний інституційний профіль та різний ступінь нормативної конкретизації, що дозволяє простежити весь спектр актів – від спеціалізованих етичних настанов до декларативних політичних заяв.

Провідну роль у регулюванні ШІ на рівні даного регіону відіграє АСЕАН, що об'єднує десять держав Південно-Східної Азії, які характеризуються безпрецедентним розмаїттям рівнів економічного розвитку та ступенів цифрової готовності. За оцінками *Government AI Readiness Index 2025* року, Сінгапур посідає 7 місце у світі у готовності до впровадження ШІ, тоді як М'янма, Лаос та Камбоджа знаходяться за межами 110-ї позиції [268]. Інституційна культура АСЕАН, що традиційно визначається як «шлях АСЕАН», ґрунтується на принципах невтручання у внутрішні справи, пріоритету діалогу та досягнення консенсусу. Як зазначає Б.А. Путра, «АСЕАН не має права приймати закони, що мають перевагу над національними політиками» [242, с. 4], що апріорі унеможлиблює створення обов'язкових нормативних актів на кшталт Акту про ШІ. За таких інституційних умов єдиним можливим форматом є акти «м'якого права».

Формування регуляторного підходу АСЕАН у сфері ШІ визначається Цифровим генеральним планом АСЕАН 2025 (англ. – *ASEAN Digital Masterplan 2025*) [241], що передбачив «розробку та прийняття регіональної політики для надання рекомендацій щодо найкращих практик управління ШІ та етики» [239, с. 7]. На його основі у лютому 2024 року АСЕАН прийняла Настанову з управління ШІ та етики (англ. – *ASEAN Guide on AI Governance and Ethics*, далі – Настанова АСЕАН) [239] – акт «м'якого права», що, як прямо визначено у його тексті, «є практичною

настановою для організацій регіону, які бажають проєктувати, розробляти та розгортати традиційні технології ШІ у комерційних та невоєнних або подвійного використання застосуваннях» [239, с. 7]. Настанова АСЕАН має добровільний характер, «не замінює та не має переваги над будь-якими чинними або майбутніми законами» [239, с. 8] і не передбачає жодних механізмів примусового виконання. Змістовна архітектура Настанови побудована навколо семи керівних принципів: прозорість та пояснюваність; справедливість та рівність; безпека та захищеність; людиноцентричність; приватність та управління даними; підзвітність та добросовісність; надійність та стійкість [239, с. 11–17] – які демонструють значний ступінь конвергенції з Принципами ОЕСР, Рекомендацією ЮНЕСКО та Етичними настановами ЄС щодо надійного ШІ, на які Настанова прямо посилається [239, с. 7], що підтверджує закономірність, ідентифіковану у Розділі 2: на рівні базових принципів спостерігається глобальний консенсус, тоді як розбіжності виявляються на рівні механізмів імплементації.

Особливої уваги заслуговує модель оцінки ризиків. На відміну від чотирирівневої класифікації Акту про ШІ, Настанова пропонує трирівневу модель визначення рівня участі людини у прийнятті рішень: «людина в контурі» (англ. – *human-in-the-loop*) – повний людський контроль; «людина над контуром» (англ. – *human-over-the-loop*) – наглядовий контроль; «людина поза контуром» (англ. – *human-out-of-the-loop*) – повністю автономні рішення ШІ [239, с. 25]. Як зазначають Я. Лу та Ф.Г. Ті, «на відміну від приписних регуляторних моделей, таких як Акт про ШІ, рамка АСЕАН робить наголос на регулюванні, заснованому на принципах, добровільному дотриманні та гнучкості, що дозволяє підприємствам та урядам приймати структури управління ШІ, адаптовані до їхніх економічних та технологічних реалій» [244, с. 404]. Ця гнучка класифікація надає організаціям повноваження самостійно визначати рівень нагляду, що є водночас і перевагою (адаптивність), і вразливістю (відсутність єдиних стандартів). У січні 2025 року АСЕАН доповнила свою регуляторну рамку Розширеною настановою з управління ШІ та етики – Генеративний ШІ [240], що стала реакцією на специфічні ризики

генеративного ШІ (створення дезінформації, дипфейків, порушення прав інтелектуальної власності). Така траєкторія – від загальної настанови до спеціалізованого доповнюючого акта зі збереженням попереднього в чинності – є проявом горизонтального інкременталізму як свідомої регулятивної відповіді на «темпоральний розрив» між швидкістю технологічних змін та повільністю нормативних процесів, обґрунтованої в підрозділі 2.3, та засвідчує реалізацію концепції «живого документа», здатного адаптуватися до технологічних змін. Загалом регуляторний підхід АСЕАН зумовлений усвідомленим вибором: для регіону з потенціалом цифрової економіки 218 млрд доларів за рік [243] стимулювання інновацій є вищим пріоритетом, ніж всеосяжне регулювання, через що АСЕАН свідомо дистанціюється від наслідування зовнішніх регуляторних моделей.

Критичний аналіз регуляторного підходу АСЕАН виявляє певні обмеження. Настанова не здатна набути переваги над національним законодавством і не передбачає санкцій за відхилення від її рекомендацій. Водночас АСЕАН свідомо дистанціюється від наслідування зовнішніх регуляторних моделей, оскільки, за словами Б.А. Путри, «не бажає нав'язувати сувору регуляторну рамку своїм державам-членам» [242, с. 4]. Ця позиція зумовлена різницею пріоритетів: для регіону з потенціалом цифрової економіки, що за оцінками досягала 218 мільярдів доларів за рік [243], стимулювання інновацій є вищим пріоритетом, ніж всеосяжне регулювання.

Принципово інший інституційний профіль демонструє Азійсько-Тихоокеанське економічне співробітництво (АТЕС) – консультативний міжурядовий форум 21 держави, рішення якого мають виключно рекомендаційний характер і не створюють юридичних зобов'язань. Тривалий час внесок АТЕС у регулювання ШІ обмежувався ситуативними згадками, зокрема доповіддю 2022 року про ШІ в розробці економічної політики [245, с. 280], проте якісний зсув відбувся у 2025 році, коли на саміті лідерів АТЕС у Кьонджу (Республіка Корея) було прийнято Ініціативу АТЕС щодо штучного інтелекту на 2026–2030 роки [273]

– перший комплексний стратегічний документ форуму, цілком присвячений проблематиці ШІ. Стратегічна логіка Ініціативи принципово відрізняється від підходу АСЕАН: якщо Настанова АСЕАН зосереджена на етичних принципах та управлінні ризиками, то Ініціатива АТЕС має переважно економіко-інфраструктурну спрямованість, побудовану навколо сприяння стійкому економічному зростанню, розбудови потенціалу ШІ та формування інвестиційної екосистеми [273, п. 2], узгоджену з ширшим порядком денним АТЕС (Бачення Путраджая 2040, Дорожня карта AIDER) [273, п. 3]. З правової точки зору Ініціатива АТЕС є актом навіть менш нормативним, ніж Настанова АСЕАН: вона не містить жодного етичного принципу, жодної класифікації ризиків і жодного механізму оцінки, обмежуючись побіжною згадкою про «безпеку, доступність, надійність» [273, п. 4(A)(iv)] без подальшого розгорнення, а натомість зосереджена на добровільних інструментах сприяння – обміні знаннями, оглядах готовності та заохоченні інвестицій приватного сектору [273, п. 4] – де множинне використання прикметника «добровільний» є промовистим маркером рівня зобов'язань. Імплементацийний механізм зводиться до інтеграції Ініціативи у діяльність форумів АТЕС з періодичними оглядами прогресу та підсумковим оглядом у 2030 році [273, п. 5].

Ще один вимір регуляторного ландшафту регіону формується Шанхайською організацією співробітництва (ШОС) – міжнародною організацією переважно безпекового та геополітичного профілю, що об'єднує десять держав-членів. У 2024–2025 роках спостерігається послідовне включення проблематики ШІ до порядку денного її вищих органів. Астанинська декларація Ради глав держав ШОС від 4 липня 2024 року стала першим актом організації, що містить розгорнуте формулювання щодо ШІ, у якому держави-члени надали «великого значення розвитку та застосуванню технологій штучного інтелекту» та зафіксували «готовність спільно працювати над запобіганням пов'язаним ризикам з метою постійного підвищення безпеки, підзвітності, надійності, довіри та справедливості технологій штучного інтелекту на благо всього людства» [271, с. 20]. Ця формула

обмежується декларативною фіксацією «готовності» без визначення конкретних зобов'язань або механізмів, проте її аксіологічне ядро – безпека, підзвітність, надійність, довіра, справедливість – демонструє конвергенцію з Принципами ОЕСР та Рекомендацією ЮНЕСКО. Перехід від декларативних заяв до інституційної інфраструктури засвідчило Спільне комюніке Ради глав урядів ШОС від 15–16 жовтня 2024 року, що зафіксувало «прийняття дорожньої карти імплементації Програми співробітництва держав-членів ШОС з розвитку штучного інтелекту» [272, с. 20]. Тяньцзінська декларація Ради глав держав ШОС від 1 вересня 2025 року суттєво поглибила цей підхід: вона артикулювала принцип суверенної рівності у сфері ШІ, підкресливши, що «всі країни мають рівні права на розвиток та використання штучного інтелекту» [270], – ця формула є концептуально значущою, оскільки вбудовує право на ШІ у ширший дискурс про рівноправну участь країн у технологічному розвитку, що принципово відрізняється як від ринково-регуляторної логіки ЄС, так і від правозахисного дискурсу Ради Європи; розширила перелік цінностей за рахунок додавання прозорості та інклюзивності порівняно з Астанинською декларацією; та зафіксувала «імплементацію Дорожньої карти виконання Програми співробітництва держав-членів ШОС з розвитку штучного інтелекту (Ченду, 12 червня 2025 р.)» [270] разом із пропозиціями щодо створення механізмів співробітництва у сфері ШІ та передових технологій.

Порівняльний аналіз актів зазначених інституцій дозволяє констатувати формування самобутньої поліцентричної моделі регулювання ШІ. АСЕАН створила спеціалізований, технічно деталізований акт, адресований безпосередньо організаціям-розробникам та користувачам, пропонуючи предметні етичні рамки з елементами оцінки ризиків; АТЕС прийняла профільну стратегічну ініціативу з конкретними механізмами імплементації, однак обмежується координуванням економічної та інфраструктурної політики без етичного чи правозахисного компонента; натомість позиція ШОС залишається інтегрованою у загальнополітичні декларації найвищого рівня, не оформленою у самостійний профільний акт. Водночас динаміка розвитку нормативної бази ШОС за період 2024–2025 років є

показовою – від одного абзацу в Астанинській декларації через прийняття дорожньої карти до розширеного формулювання в Тяньцзінській декларації з конкретними інституційними пропозиціями. Таким чином, три організації з різними інституційними профілями – АСЕАН (етичне управління), АТЕС (економічне сприяння) та ШОС (суверенна рівноправність) – паралельно формують нормативний простір на засадах добровільності. Ця модель, побудована на консенсусі та повазі до суверенітету, принципово відрізняється як від ринково-регуляторної парадигми ЄС, так і від правозахисної парадигми РЄ.

Африканський регіон демонструє принципово інший підхід до формування нормативної бази у сфері ШІ, зумовлений поєднанням амбітних цілей розвитку та фундаментальних обмежень інституційної спроможності. Інституційний контекст визначається безпрецедентним розмаїттям рівнів розвитку держав-членів – від держав з порівняно розвиненими цифровими екосистемами (Південна Африка, Кенія, Єгипет) до держав з мінімальною цифровою інфраструктурою. На відміну від ЄС та Ради Європи, чії регуляторні моделі спираються на усталені системи ринкового нагляду та правозахисного моніторингу, африканські інституції формують нормативну базу в умовах значних асиметрій інфраструктурного та інституційного розвитку. Правову основу для регулювання ШІ на рівні регіону закладає Конвенція Африканського Союзу про кібербезпеку та захист персональних даних (Конвенція Малабо), прийнята у 2014 році, яка набула чинності 8 червня 2023 року після тривалого процесу ратифікації [269] і створює нормативний фундамент для захисту даних як передумови регулювання систем ШІ. Стратегічну рамку для цифрового розвитку регіону визначила Стратегія цифрової трансформації Африки (2020–2030), що поставила за мету «трансформацію африканських суспільств та економік... з метою забезпечення інклюзивного економічного зростання, створення робочих місць, подолання цифрового розриву та викорінення бідності» [247, с. 4], включивши ШІ до переліку наскрізних тем.

Виконавча рада Африканського Союзу на своїй 45-й черговій сесії 18–19 липня 2024 року в Аккрі (Гана) схвалила Континентальну стратегію штучного

інтелекту (далі – Стратегія АС) [246], яка, як зазначено у її Передмові, «надає пріоритет розробці та адаптації систем ІІІ до африканського контексту», керуючись принципами «етики, інклюзивності та різноманітності, прав людини та людської гідності, добробуту людей, миру та процвітання» [246, с. 1]. За юридичною природою Стратегія АС є актом «м'якого права» стратегічного характеру, що споріднює її з Настановою АСЕАН з точки зору юридичної обов'язковості, проте суттєво відрізняє за змістовною парадигмою: якщо Настанова АСЕАН орієнтована на бізнес-практики та інтегрованість, Стратегія АС є передусім документом розвиткової спрямованості, що прагне вбудувати ІІІ у ширший контекст соціально-економічної трансформації континенту. Її візія – «процвітаюча та інтегрована Африка, де відповідальний, етичний та афроцентричний ІІІ є віссю інклюзивного зростання, стійкості, соціально-економічного розвитку, розширення можливостей людей та позиціонування континенту як ключового гравця у глобальному ландшафті ІІІ» [246, с. 27], а місія – використання ІІІ для прискорення соціально-економічної трансформації відповідно до Порядку денного АС 2063 та Цілей сталого розвитку поряд з мінімізацією ризиків та прискоренням розвитку потенціалу ІІІ держав-членів [246, с. 27]. Стратегія АС побудована навколо п'яти фокусних напрямів (максимізація переваг ІІІ для соціально-економічного розвитку та культурного відродження; мінімізація ризиків та захист від шкоди; нарощування потенціалу в інфраструктурі, даних, навичках та інноваціях; сприяння регіональному та міжнародному співробітництву; прискорення інвестицій у ІІІ) та п'ятнадцяти напрямів дій. Серед останніх окремої уваги заслуговують вимоги «впровадити надійне управління ІІІ, регулювання, стандарти, кодекси поведінки та найкращі практики для управління ризиками ІІІ та сприяння його зростанню» та «прийняти та впровадити етичні принципи ІІІ, що поважають права людини, гендерну рівність та гідність африканських людей, поважають різноманітність, інклюзивність та африканську культуру і цінності» [246, с. 29]. Ці вимоги важливі з двох причин: перша артикулює намір АС стати активним нормотворцем у сфері ІІІ – через впровадження регулювання, стандартів та кодексів поведінки, а не лише через

рецепцію зовнішніх рамок; друга прямо інтегрує універсальні правозахисні принципи з афроцентричною ціннісною парадигмою, що закладає нормативний фундамент для філософії *Ubuntu*, аналізованої нижче. Стратегія передбачає п'ятирічний план імплементації на 2025–2030 роки з двома фазами – створенням структур управління (2025–2026) та реалізацією основних проєктів (з 2028 року).

Стратегія АС функціонує в контексті розгалуженої багаторівневої архітектури, що охоплює спеціалізовані органи АС та регіональні економічні співтовариства. На рівні правозахисних механізмів ключову роль відіграє Резолюція 473 Африканської комісії з прав людини і народів (2021), що закликає держави-члени забезпечити сумісність розробки та використання ШІ з правами, закріпленими в Африканській хартії прав людини і народів, та наполягає на кодифікації вимоги людського контролю над ШІ як принципу прав людини [259]. На рівні економічної інтеграції паралельно діють: Альянс «Розумна Африка» (42 держави), що визначив у Програмному документі «ШІ для Африки» (2021) п'ять стовпів національних стратегій ШІ – людський капітал, перехід від лабораторії до ринку, інфраструктура, мережева взаємодія та регулювання [258]; ECOWAS з обов'язковим Додатковим актом про захист персональних даних (2010) [264]; SADC зі Стратегією цифрової трансформації (2023) [265] та ЕАС зі Стратегією цифрової трансформації на 2024–2030 роки та Рамкою управління даними (2024) [274]. Важливим етапом консолідації цих зусиль стала Африканська декларація про штучний інтелект, підписана 4 квітня 2025 року на Глобальному саміті ШІ в Кігалі представниками 54 африканських країн, АС та Альянсу «Розумна Африка», що проголошує намір створення Африканського фонду ШІ обсягом 60 мільярдів доларів та заснування Африканської ради зі штучного інтелекту [262]. Ця багатошаровість інституційної структури, хоча й свідчить про широке усвідомлення значущості проблематики ШІ, одночасно створює ризик фрагментації та дублювання зусиль в умовах обмежених інституційних ресурсів.

Однією з найбільш самобутніх рис підходу Африканського Союзу є його ціннісний фундамент, що ґрунтується на філософії *Ubuntu*. Як зазначає У.Ч.

Аджузюгу, Стратегія АС «закладає управління ІІІ в філософію *Ubuntu*, яка робить наголос на колективній користі для спільноти на протипагу індівідуальній оптимізації, на взаємозв'язку на протипагу атомістичній раціональності, та на гідності на протипагу ефективності» [249, с. 26]. Це трансформується у конкретні положення Стратегії щодо підтримки більш ніж 2000 африканських мов, захисту корінних знань, збереження культурної спадщини та уваги до колективних впливів ІІІ. Цей ціннісний підхід не є цілковитим розривом із міжнародним правом прав людини, а радше його специфічною африканською інтерпретацією: як зазначає П. Плантінга, Африканська хартія прав людини і народів «є самобутньою серед міжнародних договорів завдяки включенню колективних прав або прав солідарності «третього покоління» на додаток до індівідуальних прав» і «забезпечує потенційну основу для більш реляційної форми управління ІІІ» [260, с. 4], а сама філософія *Ubuntu* спирається на вже закріплену в африканському міжнародному праві традицію колективних прав, надаючи їй нового технологічного виміру. Це відрізняє африканський підхід від переважно індівідуалістичної правозахисної парадигми РС та ринково-регуляторної парадигми ЄС, хоча водночас К. Їлма застерігає, що нормативний зміст *Ubuntu* залишається «не цілком чітким» і потребує подальшої конкретизації для перетворення на дієві принципи регулювання ІІІ [257, с. 5]. Іншою самобутньою рисою Стратегії АС є її антиколоніальний вимір – вона «безпосередньо адресує занепокоєння щодо «ІІІ-колоніалізму», передбачаючи необхідність забезпечення захисту від експлуатації африканських працівників (модераторів контенту, анотаторів даних), які заробляють лише 2 долари на годину для іноземних компаній зі ІІІ» [249, с. 26], – ця проблематика не має аналога ні в європейських, ні в азійських регіональних актах і відображає унікальний постколоніальний контекст африканського континенту.

Водночас критичний аналіз виявляє протипроччя, яке У.Ч. Аджузіогу визначає як «парадокс розвитку-управління» (англ. – *development-governance paradox*): «напруженість між витонченими рамками управління, призначеними для зрілих технологічних екосистем, та фундаментальними обмеженнями інфраструктури,

потенціалу та ресурсів, характерними для економік, що розвиваються» [249, с. 12]. Емпіричні дані імплементації Стратегії засвідчують гостроту цього парадоксу: 83% фінансування ШІ сконцентровано лише у чотирьох країнах регіону, національні ради з ШІ існують лише у 7 з 55 держав-членів, а Регіональна незалежна рада з етики ШІ не має ані інституційної бази, ані бюджету [249, с. 31]. На відміну від ЄС, де правове регулювання ШІ побудоване на зрілій правовій та технологічній інфраструктурі, Африканський Союз стикається з необхідністю одночасного формування і правового регулювання, і базової інфраструктури. Це зумовлює унікальну модель регулювання ШІ, спрямовану на пріоритет розвитку, що, як зауважує М.Б. Юнвер, «виявляє засновану на принципах рамку, яка може бути адаптована державами-членами відповідно до їхніх спроможностей, законодавчих рамок та стратегій» [248, с. 14], та передбачає широку дискрецію у перенесенні Стратегії на національний рівень.

Американський регіон демонструє специфічну модель регіонального підходу до регулювання ШІ, що характеризується парадоксальним поєднанням інституційної насиченості та нормативної розрідженості. Аналіз цього регіону потребує послідовного дослідження діяльності трьох інституцій з принципово відмінною юридичною природою – Організації американських держав (ОАД), Економічної комісії для Латинської Америки та Карибського басейну (ЕКЛАК) та Південноамериканського спільного ринку (МЕРКОСУР).

Організація американських держав (ОАД) – регіональна міжнародна організація, заснована у 1948 році, що об'єднує 35 держав-членів і має власну міжнародну правосуб'єктність, – є ключовою інституційною платформою регіону, здатною продукувати акти з найвищим нормативним потенціалом. Резолюція Генеральної Асамблеї ОАД AG/RES. 3029 (LIV-O/24) «Зміцнення демократії» (червень 2024 року) закликає держави-члени «підтримувати або приймати ... національні стратегії для сприяння безпечній, захищеній та надійній розробці та використанню технологій ШІ у спосіб, що поважає права людини та узгоджується з міжамериканськими інструментами захисту прав людини, Рекомендацією

ЮНЕСКО щодо етики штучного інтелекту та іншими відповідними інструментами» [256, п. 2]. Ця норма є аналітично важливою, оскільки вбудовує проблематику ШІ в існуючу обов'язкову правову інфраструктуру Міжамериканської системи прав людини – Американську конвенцію з прав людини (1969), Протокол Сан-Сальвадора (1989) та юрисдикцію Міжамериканського суду з прав людини – створюючи нормативний «місток» між універсальними принципами щодо ШІ та регіональним правовим порядком. Як зауважує М.П. Лоренс, «Міжамериканський суд з прав людини, здійснюючи тлумачення договорів про права людини, має потенціал визначати конкретні обов'язки та зобов'язання для держав» у контексті регулювання ШІ, оскільки «міжнародне право прав людини надає потужні інструменти для вирішення викликів, породжених ШІ» [261, с. 93].

Декларація та План дій VII Зустрічі Міністрів з питань науки та технології ОАД (VII REMCYT, грудень 2024 року) [255] є найбільш комплексним актом ОАД у сфері ШІ: вона зобов'язує «визначити та впровадити Міжамериканський порядок денний у сфері ШІ» [255, п. 2] та рекомендує державам розробляти модельні рамки управління ШІ, адаптовані до національних контекстів [255, п. 3]. Принципово важливим є інституційне нововведення – створення двох спеціалізованих робочих груп у рамках Міжамериканського комітету з науки та технології (COMCYT): з питань моделей управління та регіонального співробітництва [255, п. 14], що свідчить про перехід від декларативного до інституційного етапу регулювання – подібно до еволюції АСЕАН від Настанов 2024 року до Розширених настанов щодо генеративного ШІ 2025 року. Водночас між підходами ОАД та АСЕАН існує принципова відмінність у адресатах: якщо АСЕАН звертається безпосередньо до приватного сектору (розробників та операторів), охоплюючи весь життєвий цикл ШІ-систем, то Декларація REMCYT зорієнтована виключно на урядову та міжурядову взаємодію, обмежується стратегічними пріоритетами без операціональних рекомендацій для окремих стадій, що зводить її практичний вплив на ринкових гравців.

Поряд з ОАД, значну роль у формуванні регіонального дискурсу відіграє Економічна комісія для Латинської Америки та Карибського басейну (ЕКЛАК). На відміну від ОАД, ЕКЛАК є однією з п'яти регіональних комісій Економічної та Соціальної Ради ООН (ЕКОСОР), тобто допоміжним органом ООН, заснованим у 1948 році, акти якого виступають формами реалізації технічного мандату і не мають самостійної нормативної сили, характерної для актів міжнародних організацій із власною правосуб'єктністю. Ключовим актом, прийнятим під егідою ЕКЛАК, є Сантьягська декларація Дев'ятої Міністерської конференції з інформаційного суспільства (7–8 листопада 2024 року) [250] – політична декларація міністерського рівня, що не породжує юридично обов'язкових зобов'язань і фіксує прагнення країн регіону «сприяти належному управлінню екосистемами штучного інтелекту... та їхнє етичне та відповідальне використання» [250, с. 1]. Водночас як слушно зауважує М.П. Лоренс, латиноамериканські стратегії «підкреслюють, що «захист або повага до прав людини» є основними цінностями при визначенні політики ШІ», проте «їм бракує визначень, що пояснюють, що «захист прав людини» означає у контексті управління ШІ» [261, с. 89]. Це різко контрастує з підходом Ради Європи, адже Рамкова конвенція про ШІ закріплює конкретні зобов'язання щодо оцінки впливу на права людини (ст. 15) та засобів правового захисту (ст. 14), перетворюючи правозахисну риторика на операціоналізовані правові механізми. На тій самій Конференції було ухвалено Цифровий порядок денний для Латинської Америки та Карибського басейну (eLAC2026) [251], який закріплює стратегічну ціль розроблення «етичних та відповідальних рамок для штучного інтелекту», що мають пріоритетувати розвиток, демократію, права людини, справедливість, верховенство права, приватність та прозорість, водночас заохочуючи інновації та регуляторну гнучкість [251, с. 5]. Зміст цих цілей подібний до принципів Настанови АСЕАН, проте з однією суттєвою різницею: якщо АСЕАН адресує свої рекомендації безпосередньо розробникам та операторам систем ШІ, то eLAC2026 залишається на рівні стратегічних цілей, адресованих урядам. Це підтверджує переважно програмний, а не регуляторний характер актів ЕКЛАК, попри передбачений

інституційний механізм моніторингу – Обсерваторію цифрового розвитку та Лабораторію політики цифрової трансформації [251, с. 8–9].

З юридичної точки зору, між актами ОАД та ЕКЛАК існує суттєва типологічна різниця: Резолюція AG/RES. 3029 та Декларація VII REMCYT є актами органів повноправної міжнародної організації, ухваленими в межах її статутних повноважень, тоді як Сантьягська декларація є рекомендаційним актом технічного секретаріату допоміжного органу ООН. Хоча обидві категорії актів належать до м'якого права, їхня потенційна нормативна траєкторія є різною: акти ОАД можуть бути враховані в авторитетному тлумаченні Американської конвенції з прав людини Міжамериканським судом з прав людини – як у консультативних висновках за ст. 64 Конвенції, так і у рішеннях за спірною юрисдикцією за ст. 62, обов'язкових для держав-учасниць, – тоді як акти ЕКЛАК інституційно обмежені рамками рекомендаційно-технічного мандату. Дослідник Е. Леві Єяті описує структурний контекст американського регіонального підходу як «трилему ШІ» – одночасне балансування «захисту прав і безпеки у нерівних суспільствах», «інновацій та конкурентоспроможності за обмежених ресурсів» і «суверенітету, відкритості та регіональної інтеграції» [263, с. 9]. Ця трилема пояснює принципову відмінність від європейського підходу: якщо ЄС, маючи потужний внутрішній ринок та розвинені інституції, може дозволити собі ресурсоємну ризикоорієнтовану модель, то у Латинській Америці та Карибському басейні – економічному ядрі регіону за межами США та Канади, де «67 відсотків робочої сили зайнято на малих та середніх підприємствах» [263, с. 11], – відтворення такої моделі створило б непропорційне навантаження на сектор малого та середнього бізнесу, здатне витіснити з ринку саме тих суб'єктів, яких політика ШІ мала б захищати.

Водночас варто зазначити, що інша впливова регіональна інтеграційна структура – Південноамериканський спільний ринок (МЕРКОСУР) досі не ухвалила жодного спеціалізованого акту з питань ШІ. Хоча Робоча група з цифрового порядку денного МЕРКОСУР проводить семінари з питань управління даними та ШІ, а Спільне комюніке президентів МЕРКОСУР у грудні 2025 року підтвердило

стратегічний характер цифрового порядку денного, ці ініціативи залишаються на доінституційному етапі. Відсутність формалізованих актів МЕРКОСУР контрастує з активністю Африканського Союзу, який за порівнянний період спромігся ухвалити розгорнуту Континентальну стратегію ШІ, і пояснюється, вірогідно, тим, що країни МЕРКОСУР (передусім Бразилія) зосередили зусилля на національному законодавстві, а не на регіональній координації.

Це дозволяє констатувати, що Американський регіон характеризується парадоксальним поєднанням інституційної насиченості (ОАД, ЕКЛАК, МЕРКОСУР, *ad hoc* саміти) та нормативної розрідженості: жоден із прийнятих актів не виходить за межі «м'якого права», не пропонує механізму класифікації ризиків чи обов'язкового оцінювання впливу на права людини, не адресований безпосередньо розробникам ШІ-систем. У порівняльній перспективі акти цього регіону перебувають на більш ранньому етапі нормативного розвитку, ніж Настанови АСЕАН чи Континентальна стратегія АС. Проте унікальною перевагою регіону є наявність розвиненої Міжамериканської системи захисту прав людини з юрисдикційним механізмом, що створює інституційні передумови для переходу від декларативної до правозастосовної моделі регулювання ШІ – потенціал, якого не мають ані АСЕАН, ані АС.

Арабський регіон представляє ще одну самотню модель, що формується на перетині амбітних національних стратегій, ісламських правових та етичних цінностей та прагнення до технологічного лідерства, та характеризується паралельною діяльністю двох організацій: Ради співробітництва арабських держав Перської затоки (РСАДПЗ), що об'єднує шість монархій Затоки, та Арабської організації з питань освіти, культури і науки (АЛЕКСО) при Лізі арабських держав, що охоплює 22 арабські держави.

На рівні РСАДПЗ ключовим актом є Настановний посібник з етики використання штучного інтелекту для держав-членів РСАДПЗ, прийнятий у листопаді 2023 року [252]. За юридичною природою це акт «м'якого права» рекомендаційного характеру, побудований навколо чотирьох базових цінностей –

людської гідності, свободи та автономії, поваги до ісламського шаріату і єдності Затоки, захисту довкілля та сталого розвитку, мирного використання для добробуту громадян [252, с. 5]. На їх основі Посібник формулює сім принципів регулювання ШІ: прийняття рішень людиною, безпека та запобігання шкоді, справедливість та недискримінація, приватність та захист даних, прозорість і пояснюваність, відповідальність та підзвітність, добросовісність та нефальсифікація [252, с. 6]. Унікальною рисою Посібника РСАДПЗ, що відрізняє його від усіх інших проаналізованих регіональних актів, є пряме включення релігійних цінностей – повага до ісламського шаріату – як одного з фундаментів етичного управління ШІ. Посібник також містить рекомендацію щодо його врахування при розробці нових політик, законодавства або національних стратегій у сферах електронного урядування, цифрової трансформації, цифрової економіки та інформаційної безпеки [252, с. 14].

Включення ісламських правових цінностей до регіональних актів у сфері ШІ не є декоративним елементом, а відображає глибоку наукову дискусію про можливість формування автономної ісламської етики ШІ. Е. Ельмахджуб обґрунтовує необхідність «плюралістичного підходу до етичного бенчмаркінгу ШІ» [266, с. 1], наголошуючи, що ісламська етична традиція містить розвинені інструменти нормативного аналізу, зокрема концепцію *масляга* (суспільного блага), яка може слугувати критерієм оцінки етичності систем ШІ; Ф. Алі, К. Бузубаа, Ф. Гелі, Б. Гамзі та С. Хан подібним чином пропонують використання етичної рамки *i'timani* (відповідальність за доручене) як «уніфікованої етичної моделі для технологій ШІ», заснованої на концепції божественної довіри [267, с. 1]. Ці наукові розробки засвідчують, що посилення Посібника РСАДПЗ на ісламський шаріат є не просто культурним маркером, а входом у розвинену альтернативну аксіологічну систему, що споріднює підхід РСАДПЗ з підходом Африканського Союзу через апеляцію до власних незахідних ціннісних систем та відрізняє обидва від суто секулярних підходів Азійсько-Тихоокеанського та Американського регіонів.

Водночас Арабський регіон представлений не лише РСАДПЗ, а й діяльністю структур Ліги арабських держав (ЛАД), яка хоча й не прийняла спеціалізованих актів щодо ШІ, розвиває цю проблематику через свою спеціалізовану установу – Арабську організацію з питань освіти, культури і науки (АЛЕКСО). У червні 2025 року АЛЕКСО прийняла Хартію з етики штучного інтелекту (англ. – *ALECSO Charter on Artificial Intelligence Ethics*) [254] – секторальний акт, спрямований на регулювання використання ШІ у трьох сферах компетенції організації: освіті, культурі та науці. Хартія позиціонується як «довідкова рамка для спрямування етичних та практичних політик щодо цих технологій в арабському світі» [254, с. 4] і формулює цінності, серед яких – повага до людської гідності та прав людини, справедливість та соціальна рівність, відповідальні інновації, сприяння сталому розвитку, різноманітність та інклюзивність, збереження арабської культурної ідентичності та просування арабської спадщини. За сферою застосування Хартія адресована «всім державам-членам АЛЕКСО, включаючи уряди, відповідальні за встановлення правових та регуляторних рамок для систем ШІ протягом їхнього життєвого циклу, та активним інституціям у цій сфері» [254, с. 47].

Порівняльний аналіз Хартії АЛЕКСО та Посібника РСАДПЗ виявляє водночас і спільні риси, і суттєві відмінності: обидва акти є «м'яким правом» рекомендаційного характеру, апелюють до ісламських та арабських культурних цінностей і підкреслюють збереження культурної ідентичності, проте Посібник РСАДПЗ орієнтований на загальне етичне регулювання ШІ у шести державах Затоки, тоді як Хартія АЛЕКСО має виражений секторальний характер (освіта, культура, наука) та ширшу географію – 22 арабські держави з істотно різними рівнями технологічного розвитку. Як слушно зазначають М.Р. Албусом та О.Р. Аль-Джайюсі, регулювання «ШІ у РСАДПЗ здійснюється переважно через «м'яке» регулювання, яке надає перевагу стратегіям та етичним настановам перед обов'язковими правилами», проте «стратегії РСАДПЗ посиляються на етику та культурну чутливість» і «залишаються переважно декларативними без надійних

регуляцій та залучення зацікавлених сторін» [253, с. 2414], що піднімає питання щодо ефективного застосування етичних принципів.

Здійснений аналіз чотирьох позаєвропейських регіонів у поєднанні з результатами аналізу правової бази ЄС та діяльності Ради Європи дозволяє виявити ключові закономірності формування міжнародно-правового регулювання ІІІ на регіональному рівні. Це порівняння систематизовано у нижченаведеній таблиці 3.

Таблиця 3. Порівняльна характеристика регіональних підходів до регулювання

ІІІ

Організація/об'єднання	Ключовий акт	Юридична природа	Обов'язковість	Основні адресати	Модель регулювання	Ціннісний фундамент
ЄС	Регламент (ЄС) 2024/1689 (Акт про ІІІ)	«Тверде право»: регламент прямої дії	Обов'язковий для всіх держав-членів та суб'єктів ринку	Розробники, постачальники, оператори ІІІ, наглядові органи	Ризикоорієнтована (4 рівні ризику)	Основоположні права ЄС, безпека продукції, конкурентоспроможність
Рада Європи	Рамкова конвенція про ІІІ та права людини, демократію і верховенство права (2024)	«Тверде право»: юридично обов'язковий міжнародний договір	Обов'язковий після ратифікації	Держави-учасниці	Правозахисна рамкова (принципи + звітність)	Права людини, демократія, верховенство права (ЄКПЛ)
АСЕАН	Настанова з управління ІІІ та етики (2024); Розширена настанова — Генеративний ІІІ (2025)	«М'яке право»: добровільна настанова	Добровільний	Організації-розробники та користувачі ІІІ	Заснована на принципах, елементи оцінки ризиків (3 рівні участі людини)	Інтероперабельність, стимулювання цифрової економіки
АТЕС	Ініціатива АТЕС щодо ІІІ на 2026–2030 (2025)	«М'яке право»: стратегічна ініціатива	Добровільний	Уряди держав-учасниць	Економіко-інфраструктурна (сприяння розвитку)	Економічне зростання, інвестиції, інфраструктура
ШОС	Астанинська (2024) та Тяньцзіньська (2025)	«М'яке право»: політичні декларації	Декларативний	Держави-члени	Безпеково-суверенна (декларативна)	Суверенна рівноправність, безпека

	декларації; Дорожня карта					
АС	Континентальна стратегія ШІ (2024)	«М'яке право»: стратегічний документ	Добровільний, потребує імплементації	Держави-члени	Розвиткова (пріоритет нарощування потенціалу)	Філософія Ubuntu, антиколоніальний вимір
ОАД	Резолюція AG/RES. 3029 (2024); Декларація VII REMCUT (2024)	«М'яке право»: резолюції та декларації	Рекомендаційний	Держави-члени	Правозахисна декларативна	Міжамериканська система прав людини, подолання нерівності
ЕКЛАК	Сантьягська декларація (2024); eLAC2026	«М'яке право»: політична декларація та стратегія	Рекомендаційний	Уряди	Програмно-стратегічна	Інклюзивний розвиток, подолання цифрового розриву
РСАДПЗ	Настановний посібник з етики ШІ (2023)	«М'яке право»: рекомендаційний посібник	Рекомендаційний	Уряди	Етично-рекомендаційна	Ісламський шаріат, людська гідність
АЛЕКСО	Хартія з етики ШІ (2025)	«М'яке право»: секторальна хартія	Рекомендаційний	Уряди та інституції арабських держав	Секторальна етична (освіта, культура, наука)	Арабська культурна ідентичність, ісламські цінності

Порівняльний аналіз дозволяє виявити п'ять ключових закономірностей. Найбільш очевидною є спектр юридичної обов'язковості, що охоплює діапазон від «твердого права» (Акт про ШІ, Конвенція РЄ) до різноманітних форм «м'якого права» (Настанова АСЕАН, Стратегія АС, посібник РСАДПЗ, секторальна хартія АЛЕКСО): жодна регіональна організація поза Європою не створила юридично обов'язкового акту, тож «тверде право» у сфері регулювання ШІ залишається на сьогодні виключно європейським феноменом. Типологічна різноманітність форм регулювання детермінована інституційною архітектурою відповідної організації – наднаціональна інтеграційна модель ЄС логічно породжує обов'язковий горизонтальний акт з санкційним механізмом, конвенційна правоохоронна модель Ради Європи – рамковий договір з механізмом звітності, «шлях АСЕАН» – добровільну настанову, розвитковий мандат АС – стратегічний документ, культурно-освітній мандат АЛЕКСО – секторальну хартію, – що свідчить про

предетермінованість регіональних підходів до ШІ попереднім інституційним досвідом кожної організації.

Друга закономірність – аксіологічний консенсус при регуляторному розмаїтті: усі без винятку проаналізовані акти розділяють базові принципи прозорості, справедливості, підзвітності, безпеки та людиноцентричності, що відображає нормативний вплив Рекомендації ЮНЕСКО та Принципів ОЕСР щодо ШІ, на які прямо посилаються Настанова АСЕАН [239], Стратегія АС [246], Хартія АЛЕКСО [254] та Сантьягська декларація. Водночас кожен регіон привносить у глобальний дискурс власні ціннісні акценти: Азійсько-Тихоокеанський – орієнтацію на практичну інтеперабельність та стимулювання цифрової економіки; Африканський – філософію *Ubuntu*, антиколоніальний вимір та пріоритет розвитку; Американський – акцент на подоланні структурної нерівності та потенціал Міжамериканської системи захисту прав людини; Арабський – інтеграцію ісламських етичних цінностей. Фундаментальні ж розбіжності виявляються на рівні механізмів імплементації, де одні й ті самі принципи втілюються у штрафи та аудити (ЄС), рамкові зобов'язання (РЄ), корпоративну саморегуляцію (АСЕАН), стратегічне планування (АС), політичні декларації або етичні посібники та хартії (РСАДПЗ, АЛЕКСО).

Третя закономірність пов'язана з диференційованим впливом «Брюссельського ефекту». Екстратериторіальний обсяг дії Акту про ШІ поширюється на всі регіони, проте цей вплив не є однорідним: в Азійсько-Тихоокеанському регіоні він створює парадоксальну ситуацію, коли *de jure* регулювання залишається добровільним, тоді як *de facto* компанії, орієнтовані на глобальні ринки, змушені дотримуватися норм права ЄС; у Латинській Америці – проявляється переважно через імітацію ризикоорієнтованого підходу у національних законопроектах Бразилії, Чилі та Колумбії, тоді як на регіональному рівні організації послідовно апелюють до принципів ОЕСР та Рекомендації ЮНЕСКО; в АСЕАН та Африканському Союзі держави свідомо відмежовуються від європейської моделі; в арабському регіоні вплив є мінімальним, оскільки ані

Посібник РСАДПЗ, ані Хартія АЛЕКСО не відтворюють положення Акту про ШІ, апелюючи до власних культурно-релігійних ціннісних систем. Показово, що екстратериторіальний тиск «Брюссельського ефекту» стимулює поступовий рух окремих держав від «м'якого права» на регіональному рівні до «твердого права» на національному – як зазначають Ц. Сьюй, Т. Лі та Дж. Гоггін, «тенденція є очевидною: азійські країни рухаються до «твердого регулювання» через встановлення правил та законів щодо ШІ» [245, с. 280] на національному рівні, що підтверджується прийняттям обов'язкових законів про ШІ у Республіці Корея та Японії, та засвідчує, що «Брюссельський ефект» Акту про ШІ має реальний, але обмежений вплив і не здатний встановити єдиний глобальний підхід до регулювання ШІ.

Четверта закономірність виявляє парадокс глобального регулювання ШІ: регіони з найбільшими потребами у регулюванні мають найменші інституційні та фінансові спроможності для створення ефективних регуляторних рамок, натомість регіон із найрозвиненішою регуляторною рамкою (ЄС) не є основним центром розробки систем ШІ – ця роль належить США та Китаю, які, попри членство в окремих регіональних організаціях, не є учасниками жодного юридично обов'язкового регіонального акту щодо ШІ та розвивають переважно національні регуляторні моделі.

Нарешті, п'ята закономірність стосується формування «регуляторного полілогу», в якому різні регіони артикулюють власні нормативні позиції, що відображають їхні специфічні потреби, цінності та інституційні можливості, замість поширення єдиного регуляторного стандарту. Цей полілог збагачує глобальний дискурс перспективами, традиційно периферійними, – регіональними ціннісними акцентами, які привносять у міжнародно-правовий дискурс виміри, не репрезентовані в європейських моделях, – та одночасно створює як можливості для формування більш інклюзивного глобального регулювання, так і ризики фрагментації регуляторного простору та «регуляторного арбітражу».

Для України порівняльний аналіз регіональних підходів має практичне значення у трьох вимірах: підтверджує стратегічну обґрунтованість обраного курсу

на європейську модель регулювання як таку, що забезпечує юридично обов'язковий рівень захисту прав людини у контексті ШІ; дозволяє врахувати досвід інших регіонів – зокрема, гнучкість АСЕАН у стимулюванні інновацій та розвитковий акцент Африканського Союзу – при формуванні збалансованого національного законодавства; а виявлений у контексті АС «парадокс розвитку-управління» є релевантним і для України, яка в умовах повоєнної відбудови стикатиметься з необхідністю одночасного формування базової інфраструктури ШІ та створення належних регуляторних рамок – завдання, що вимагатиме не лише нормативного наслідування європейського зразка, а й творчої адаптації найкращих практик різних регіональних моделей.

Висновки до Розділу 3

Проведене комплексне дослідження міжнародно-правового регулювання ШІ на регіональному рівні дозволяє сформулювати такі узагальнюючі висновки щодо поточного стану, закономірностей і тенденцій розвитку цього регулювання, а також його співвідношення з універсальним та міжрегіональним рівнями.

Формування правової бази ЄС у сфері регулювання ШІ охоплює п'ять фаз: від фази парламентської ініціативи (2015–2017) через стратегічне планування (2018), формування етико-правового фундаменту (2019) та перехід до юридично обов'язкового регулювання (2020–2024) – до сучасної фази масштабних стратегічних ініціатив (2025 і далі). Запропонована періодизація виявляє три системні закономірності підходу ЄС: дуалізм ринкового та ціннісного обґрунтування регулювання, що створює іманентну напруженість між цілями захисту прав та стимулювання інновацій; поступове послаблення механізмів забезпечення відповідності під впливом лобістських зусиль транснаціональних корпорацій технологічного сектору, що проявляється у запровадженні самооцінки та самокласифікації ризиків замість незалежної верифікації; а також реактивний характер регулювання, що наочно продемонструвала поява генеративного ШІ.

Регламент (ЄС) 2024/1689 (Акт про ШІ) [25], попри його безпрецедентний характер як першого юридично обов'язкового горизонтального акта щодо ШІ, що набув чинності, має структурні обмеження, зумовлені обранням ст. 114 Договору про функціонування ЄС (ДФЄС) як правової основи, що позиціонує регулювання ШІ передусім як інструмент гармонізації законодавства держав-членів у межах внутрішнього ринку, а не як самостійний правозахисний механізм.

Порівняння нульового проєкту Рамкової конвенції Ради Європи про штучний інтелект та права людини, демократію і верховенство права (CAI(2023)01) [225] з її фінальним текстом 2024 року [24] засвідчує тенденцію до систематичного звуження обсягу зобов'язань у ході міжурядових переговорів за чотирма ключовими параметрами: відхід від прямої застосовності до суб'єктів приватного права, вилучення екологічного виміру, усунення прямого посилання на юридичну відповідальність та розширення винятку щодо національної безпеки. Ця тенденція є структурно аналогічною тій, що простежується в еволюції Акту про ШІ, що дає підстави для висновку про загальну закономірність: на етапі переговорів правозахисний потенціал юридично обов'язкових актів щодо регулювання ШІ послідовно зменшується під впливом як державних, так і корпоративних інтересів. Водночас ці два акти різної правової природи можуть бути взаємодоповнюваними: Рамкова конвенція реалізує правозахисний підхід, тоді як Акт про ШІ – ризикоорієнтований підхід.

Конвенція про захист прав людини і основоположних свобод 1950 року (ЄКПЛ) демонструє значний адаптивний потенціал щодо проблематики ШІ, що підтверджує практика Європейського суду з прав людини (ЄСПЛ): Суд послідовно розширює тлумачення конвенційних прав для охоплення ситуацій, пов'язаних із застосуванням цифрових технологій, формуючи прецедентну базу, яка поступово пристосовується до реалій цифрової доби. Разом з тим ця практика має переважно реактивний характер і не може замінити проактивне правове регулювання ШІ, що здійснюється через Рамкову конвенцію та Акт про ШІ.

Порівняльний аналіз європейських та позаєвропейських регіональних ініціатив виявляє п'ять ключових закономірностей міжнародно-правового регулювання ШІ на регіональному рівні. Перша – «спектр юридичної обов'язковості»: «тверде право» у сфері регулювання ШІ залишається феноменом, притаманним виключно європейському регіону, а типологічна різноманітність форм регулювання зумовлена інституційною архітектурою кожної організації, що свідчить про детермінованість регіональних підходів попереднім інституційним досвідом. Друга – «аксіологічний консенсус при регуляторному розмаїтті»: усі проаналізовані акти поділяють базові принципи прозорості, справедливості, підзвітності та людиноцентричності, проте кожен регіон привносить у глобальний дискурс власні ціннісні акценти, а фундаментальні розбіжності виявляються саме на рівні механізмів імплементації. Третя – диференційований характер «Брюссельського ефекту»: ефект справляє реальний, але обмежений вплив і не здатен встановити єдиний глобальний підхід до регулювання ШІ. Четверта – «парадокс глобального регулювання ШІ»: регіони з найбільшими потребами у захисті прав людини від ризиків ШІ мають найменші інституційні та фінансові спроможності для створення ефективних регуляторних рамок. П'ята – формування «регуляторного полілогу», в якому різні регіони артикулюють власні нормативні позиції, що збагачують глобальний дискурс перспективами, не репрезентованими в європейських регуляторних інструментах, водночас створюючи ризики фрагментації регуляторного простору.

Усьому регіональному рівню регулювання притаманні стійкі спільні прогалини: імплементаційний дефіцит, особливо відчутний у регіонах Глобального Півдня; проблема інтероперабельності між регіональними режимами; регуляторне запізнення, зумовлене темпами розвитку генеративного ШІ; а також недостатня увага до впливу систем ШІ на трудові відносини та довкілля.

У підсумку сформувалася багаторівнева система міжнародно-правового регулювання ШІ, в якій кожен рівень виконує специфічну та незамінну функцію: універсальний забезпечує легітимність та інклюзивність через консенсусні

резолуції ГА ООН та етичні рамки спеціалізованих установ; міжрегіональний здійснює консолідацію експертних підходів та вироблення практично застосовних стандартів через механізми ОЕСР, G7 та G20; регіональний демонструє найвищу інституційну спроможність до переходу від декларування принципів до юридично обов'язкових норм – спроможність, наразі реалізовану лише Радою Європи через Рамкову конвенцію та ЄС через Акт про ШІ. Поточний стан цієї системи характеризується асиметрією, а основними тенденціями її розвитку є рух від добровільних принципів до обов'язкових норм, від фрагментарних ініціатив до системних регуляторних режимів та від технологічно нейтрального регулювання до специфічних норм, адресованих саме системам ШІ. Стратегічним викликом залишається забезпечення взаємоузгодженості регіональних регуляторних режимів, що потребує не механічної уніфікації моделей, а розвитку механізмів взаємного визнання та координації між ними задля запобігання фрагментації глобального регуляторного простору та збереження балансу між стимулюванням інновацій і захистом прав людини.

ВИСНОВКИ

Проведене дисертаційне дослідження дозволяє сформулювати такі висновки.

Формування концепції штучного інтелекту охоплює понад сім десятиліть – від Дартмутської конференції 1956 року, для проведення якої Дж. Маккарті запропонував сам термін «штучний інтелект», до процесу інтенсивної нормотворчої діяльності, що триває нині. У міжнародно-правовому дискурсі це формування відбувалося нелінійно: ШІ пройшов шлях від суто технічного феномену до об'єкта міжнародно-правового регулювання – через етичні принципи та інструменти «м'якого права» до ухвалення у 2024 році перших юридично обов'язкових актів у цій сфері, які, однак, не становлять кінцевої точки цього процесу, а знаменують якісний перехід до нового етапу регулювання. Систематичні спроби формування міжнародно-правового регулювання технології ШІ почалися лише в середині 2010-х років, що зумовлено сукупністю взаємопов'язаних чинників технологічного, соціально-політичного та геополітичного характеру, зокрема проривом у галузі глибокого навчання, проникненням систем ШІ у критичні сфери суспільного життя та перетворенням ШІ на стратегічну технологію, що визначає економічну конкурентоспроможність держав.

Єдине і загальноприйняте визначення ШІ відсутнє як у науковій доктрині, так і в міжнародно-правових актах. Різні міжнародні організації та міждержавні об'єднання розробили власні визначення або робочі описи ШІ, кожне з яких відображає специфічні пріоритети та правові традиції відповідної організації. На підставі аналізу міжнародно-правових актів та наукової доктрини запропоновано авторську класифікацію семи концептуальних підходів до визначення ШІ: технологічного, антропоцентричного, функціонального, системного, еволюційного, цілеорієнтованого та інтегративного. На цій основі сформульовано авторську дефініцію штучного інтелекту як комплексної соціотехнічної системи, що характеризується функціональною здатністю до інференції (виведення висновків з даних), операційною автономністю різного ступеня, адаптивністю, генеративною

спроможністю, транскордонним характером застосування, потенціалом подвійного використання та створенням юридично значущих наслідків для сфери прав людини, міжнародного правопорядку та безпеки. Запропонована дефініція втілює інтегративний підхід, забезпечуючи технологічну нейтральність та придатність для використання у міжнародно-правових актах різного рівня.

Міжнародне право не закріпило за штучним інтелектом самостійного правового статусу. На відміну від держави та міжнародної організації як загальноновизнаних суб'єктів міжнародного права, а також індивіда, для якого міжнародне право встановлює спеціально визначені статусні категорії – біженця, особи без громадянства, комбатанта – із чітким переліком прав, обов'язків та гарантій, щодо штучного інтелекту така нормативна конструкція відсутня як у чинних юридично обов'язкових актах, так і в актах «м'якого права». Натомість міжнародне право здійснює комплексне регулювання ШІ як об'єкта, покладаючи зобов'язання на своїх традиційних суб'єктів – держави та міжнародні організації; регуляторні ж вимоги до суб'єктів приватного права, які беруть участь у життєвому циклі систем ШІ, встановлюються опосередковано – через національне право держав та право Європейського Союзу.

Обґрунтовано правову природу ШІ як об'єкта міжнародно-правового регулювання *sui generis*, що характеризується поєднанням матеріальних і нематеріальних складових, об'єктних і суб'єктних рис, динамічності та еволюційності. Основними характеристиками ШІ, що детермінують особливості формування відповідного правового режиму, є: транскордонний характер, що створює виклик для традиційної системи міжнародного права, побудованої на принципі територіального суверенітету держав; потенціал подвійного використання, що проявляється в можливості застосування однієї технології як для цивільних, так і для воєнних цілей; автономність різного ступеня – від повної залежності від людських інструкцій до високого рівня самостійності в прийнятті рішень; здатність до еволюції та самовдосконалення, що породжує проблему «рухомої цілі» (англ. – *moving target*) для правового регулювання.

Формуванню міжнародно-правового регулювання ШІ притаманна фрагментація регуляторних підходів, що проявляється на всіх рівнях – універсальному, міжрегіональному та регіональному. На універсальному рівні відсутній єдиний міжнародний договір, навіть рамковий, присвячений регулюванню ШІ; замість цього спостерігається мозаїка різномірних ініціатив міжнародних організацій. Водночас простежується стійка конвергенція принципів і цінностей регулювання ШІ навколо ядра спільних засад: людиноцентричного підходу, прозорості та пояснюваності, підзвітності та людського нагляду, безпеки та надійності, інклюзивності та недискримінації, захисту приватності та даних. Ця конвергенція є не випадковою: вона відображає утвердження «принципізму» (англ. – *principism*) – підходу, що ґрунтується на формулюванні високорівневих етичних принципів як основи регулювання, – як спільної методологічної основи більшості етичних рекомендацій, правових орієнтирів та кодексів поведінки. Цей аксіологічний консенсус створює підґрунтя для формування універсальних норм міжнародного права у сфері регулювання ШІ, хоча кожен рівень регулювання та кожна організація привносять у цей процес власні акценти, зумовлені їхніми функціями та мандатами. «М'яке право» продовжує відігравати ключову роль у цьому регулюванні навіть після появи перших актів «твердого права», забезпечуючи його гнучкість та адаптивність до швидких технологічних змін.

Співвідношення «м'якого» та «твердого» права у сфері регулювання ШІ має виражену специфіку. «М'яке право» на універсальному та міжрегіональному рівнях виконує щонайменше чотири функції, які мають критичне значення для формування міжнародно-правового регулювання ШІ: функцію кристалізації консенсусу – через поступове узгодження принципів, термінології та концептуальних рамок у численних актах різних інституцій; функцію правового експерименту – тестування різних регуляторних підходів без ризиків, пов'язаних із передчасною фіксацією обов'язкових норм у технологічній сфері, що швидко змінюється; функцію легітимізації – формулювання принципів на рівні ГА ООН, ЮНЕСКО або ОЕСР створює нормативні очікування, що впливають на поведінку держав та інших

учасників міжнародних відносин навіть за відсутності юридичного примусу; функцію підготовки підґрунтя для юридично обов'язкового регулювання – принципи, апробовані у «м'якому праві», поступово інкорпорується до юридично обов'язкових актів регіонального та національного рівнів. Водночас саме лише «м'яке право» є недостатнім для ефективного регулювання технологій із масштабними та потенційно необоротними наслідками, до яких належить ШІ, що обумовлює необхідність його доповнення юридично обов'язковими актами, які формуються передусім на регіональному рівні.

Основу міжнародно-правового регулювання ШІ на універсальному рівні становлять дванадцять принципів, систематизованих за результатами аналізу 23 резолюцій Генеральної Асамблеї ООН за період 2015–2025 років: людиноцентричний підхід; безпека, захищеність та надійність; прозорість та пояснюваність; підзвітність та людський нагляд; інклюзивність та недискримінація; міжнародне співробітництво; подолання цифрового розриву; сприяння сталому розвитку; захист приватності та даних; підхід за участю багатьох зацікавлених сторін; обережність (запобіжний підхід); застосовність міжнародного права. Розвиток нормотворчої активності ГА ООН у сфері ШІ охоплює чотири періоди – від «Підготовчої фази» (2015–2022) через «Пробудження інтересу» (2023) та «Інституціоналізацію» (2024) до «Оперативної фази» (2025), – причому 70% усіх досліджуваних резолюцій (16 з 23) ухвалені протягом 2024–2025 років. Цьому розвитку притаманні три ключові закономірності: систематизація підходу на основі «життєвого циклу» систем штучного інтелекту як базової регуляторної основи, стандартизація тріади «безпечні, захищені та надійні системи штучного інтелекту» та позиціонування Глобального цифрового договору як точки конвергенції нормативного регулювання ШІ на рівні ООН.

Діяльність спеціалізованих установ ООН формує багатовимірну систему регулювання ШІ, що поєднує етичні принципи ЮНЕСКО, технічні стандарти МСЕ, галузеві керівництва ВООЗ, аналітичні дослідження ВОІВ та МОП. Ці ініціативи конкретизують загальні принципи, встановлені резолюціями ГА ООН; разом із

самими резолюціями та актами міжрегіональних ініціатив вони утворюють «градацію нормативності» на універсальному та міжрегіональному рівнях, що проявляється у вигляді «концентричних кіл»: у центрі розташовані акти з найвищим ступенем політичної легітимності і водночас найвищою абстрактністю – резолюції ГА ООН; друге коло формують акти спеціалізованих установ, які деталізують ці принципи через секторальні механізми та стандарти; третє коло складають акти міжрегіональних ініціатив, що балансують між загальністю та конкретністю. Ця градація є не лише описовою характеристикою чинної системи, але й її функціональною перевагою, оскільки різні ступені нормативності дозволяють різним категоріям учасників відносин у сфері ШІ знаходити відповідні нормативні орієнтири. Така система є «слабким режимним комплексом» (англ. – *weak regime complex*) – поліцентричним інституційним ландшафтом без чіткої ієрархії, в якому акти різного ступеня нормативності співіснують та підсилюють один одного.

У системі міжнародно-правового регулювання ШІ виокремлено самостійний міжрегіональний рівень – на підставі чотирьох авторських критеріїв: міжконтинентального складу учасників відповідних об'єднань, відсутності формального географічного обмеження їхніх мандатів, переважної орієнтації на економічне та технологічне співробітництво, домінування інструментів «м'якого права». Центральною ланкою цього рівня виступає ОЕСР, до якої тяжіють пов'язані формати G7 та G20, а також GPAI, інтегроване з 2024 року в її інституційну структуру. Міжрегіональні ініціативи виконують у поліцентричному режимному комплексі роль «м'яких консолідаторів», яка є водночас цінною, оскільки дозволяє досягати координації за відсутності універсального «твердого права», та обмеженою, оскільки не гарантує ані універсальності охоплення, ані ефективності виконання. Взаємодія між універсальним та міжрегіональним рівнями є двоспрямованою: Рекомендація ОЕСР 2019 року передувала резолюціям ГА ООН 2024–2025 років і концептуально вплинула на їх формулювання, тоді як Рекомендація ЮНЕСКО 2021 року стала етичним фундаментом, на який спираються ініціативи як спеціалізованих установ, так і міжрегіональних форматів.

Виявлена тенденція циклічного оновлення актів дістала теоретичне узагальнення в авторській концепції «ітеративного нормотворення» – моделі формування актів «м'якого права» у сфері міжнародно-правового регулювання ШІ, за якої такий акт не розглядається як остаточна фіксація норм, а функціонує як «живий документ» (англ. – living document), що проходить повторювані цикли прийняття, апробації, оцінки ефективності та перегляду у відповідь на технологічні зміни. Ця модель відрізняється від класичної парадигми міжнародного нормотворення за трьома ключовими ознаками: нормативна незавершеність як свідомий вибір, скорочення тривалості нормотворчих циклів, а також інкрементальність змін із збереженням ціннісного ядра акта. Практичним втіленням цієї моделі є оновлення Рекомендації Ради ОЕСР щодо штучного інтелекту та систематичний перегляд керівних принципів (настанов) АСЕАН з управління ШІ.

Розвитку міжнародно-правового регулювання ШІ на універсальному та міжрегіональному рівнях притаманні п'ять ключових тенденцій: прискорення нормотворчої діяльності, що проявляється у стрибкоподібному зростанні кількості актів протягом 2023–2025 років; перехід від фрагментарного до системного регулювання; інституціоналізація регулювання через створення постійних структур; розширення предмета та поглиблення деталізації регулювання – від загальних декларацій до регламентації конкретних сфер застосування ШІ; поступове формування моделі ітеративного нормотворення. Водночас зберігаються стійкі прогалини, спільні для досліджуваних рівнів: жоден з них не забезпечує адекватного регулювання таких критичних питань, як вплив технологій ШІ на довкілля; відповідальність суб'єктів приватного права за наслідки функціонування систем ШІ; виклики, пов'язані з генеративним ШІ (дипфейки, «галюцинації», порушення авторських прав); проблематика загального штучного інтелекту (ЗШІ) та довгострокових екзистенційних ризиків; воєнне використання ШІ. Ці прогалини набувають особливої гостроти на тлі «темпорального розриву» між швидкістю технологічного розвитку та повільністю нормотворчих процесів.

Саме на регіональному рівні розпочався перехід від декларування принципів до формулювання юридично обов'язкових норм. Тим самим регіональний рівень виконує функцію, яку поки що не змогли забезпечити універсальний та міжрегіональний рівні: трансформацію «м'якого права», апробованого на цих рівнях, у нормативні приписи з юридично обов'язковою силою. Формування правової бази ЄС у сфері регулювання ШІ охоплює п'ять фаз: від фази парламентської ініціативи (2015–2017) через стратегічне планування (2018), формування етико-правового фундаменту (2019) та перехід до юридично обов'язкового регулювання (2020–2024) – до сучасної фази масштабних стратегічних ініціатив (2025 і далі). Підходу ЄС притаманні три системні закономірності: дуалізм ринкового та ціннісного обґрунтування регулювання, що створює іманентну напруженість між цілями захисту прав і стимулювання інновацій; поступове послаблення механізмів забезпечення відповідності під впливом лобістських зусиль транснаціональних корпорацій технологічного сектору; реактивний характер регулювання, що наочно продемонструвала поява генеративного ШІ. Водночас Регламент (ЄС) 2024/1689 (Акт про ШІ) став першим юридично обов'язковим горизонтальним нормативним актом щодо штучного інтелекту, який набув чинності, і справляє глобальний вплив через «Брюссельський ефект» (англ. – Brussels effect).

Рамкова конвенція Ради Європи про штучний інтелект та права людини, демократію і верховенство права 2024 року (CETS № 225) та Регламент (ЄС) 2024/1689 (Акт про ШІ) є взаємодоповнюваними інструментами різної правової природи: Конвенція – міжнародний договір, відкритий для приєднання держав з різних регіонів, – реалізує правозахисний підхід, тоді як Акт про ШІ – інструмент права ЄС – ризикоорієнтований підхід. Водночас порівняння нульового проєкту Рамкової конвенції з її фінальним текстом засвідчує тенденцію до систематичного звуження обсягу зобов'язань у ході міжурядових переговорів за чотирма параметрами: відхід від прямої застосовності до суб'єктів приватного права, вилучення екологічного виміру, усунення прямого посилання на юридичну

відповідальність та розширення винятку щодо національної безпеки. Аналогічна закономірність простежується і в еволюції Акту про ШІ, що дає підстави для висновку про загальну тенденцію: на етапі переговорів правозахисний потенціал юридично обов'язкових актів щодо регулювання ШІ послідовно зменшується під впливом як державних, так і корпоративних інтересів.

Порівняльний аналіз регіональних ініціатив виявляє п'ять ключових закономірностей міжнародно-правового регулювання ШІ на регіональному рівні. Першою є «спектр юридичної обов'язковості»: «тверде право» у сфері регулювання ШІ залишається нині феноменом, притаманним виключно європейському регіону, а типологічна різноманітність форм регулювання детермінована інституційною архітектурою кожної організації. Другою – «аксіологічний консенсус при регуляторному розмаїтті»: усі проаналізовані акти поділяють базові принципи прозорості, справедливості, підзвітності та людиноцентричності, однак кожен регіон привносить у глобальний дискурс власні ціннісні акценти: Азійсько-Тихоокеанський регіон – орієнтацію на практичну інтероперабельність; Африканський – філософію Ubuntu та антиколоніальний вимір; Американський – наголос на подоланні структурної нерівності; Арабський – інтеграцію ісламських етичних концепцій. Третьою – диференційований характер «Брюссельського ефекту»: ефект справляє реальний, але обмежений вплив і не здатен встановити єдиний глобальний підхід до регулювання ШІ. Четвертою – «парадокс глобального регулювання ШІ»: регіони з найбільшими потребами у захисті прав людини від ризиків ШІ, передусім Глобального Півдня, мають найменші інституційні та фінансові спроможності для створення ефективних регуляторних рамок, що породжує асиметрію між декларованими цілями та реальними можливостями їх досягнення. П'ятою – формування «регуляторного полілогу», в якому різні регіони артикулюють власні нормативні позиції, що збагачують глобальний дискурс перспективами, не репрезентованими в європейських регуляторних інструментах. Усьому регіональному рівню притаманні також стійкі спільні прогалини: імплементаційний дефіцит, особливо відчутний у регіонах Глобального Півдня;

проблема інтероперабельності між регіональними режимами; регуляторне запізнення, зумовлене темпами розвитку генеративного ШІ; а також недостатня увага до впливу систем ШІ на трудові відносини та довкілля.

У підсумку сформувалася багаторівнева система міжнародно-правового регулювання ШІ, в якій кожен рівень виконує специфічну та незамінну функцію: універсальний забезпечує легітимність та інклюзивність через консенсусні резолюції ГА ООН та етичні рамки спеціалізованих установ; міжрегіональний здійснює консолідацію експертних підходів та вироблення практично застосовних стандартів через механізми ОЕСР, G7 та G20; регіональний демонструє найвищу інституційну спроможність до переходу від декларування принципів до юридично обов'язкових норм – спроможність, наразі реалізовану лише Радою Європи через Рамкову конвенцію та ЄС через Акт про ШІ. Поточний стан цієї системи характеризується асиметрією та домінуванням «м'якого права» як нормативного фундаменту: жоден з проаналізованих актів універсального та міжрегіонального рівнів не має юридично обов'язкової сили. Основними тенденціями розвитку є рух від добровільних принципів до обов'язкових норм, від фрагментарних ініціатив до системних регуляторних режимів та від технологічно нейтрального регулювання до специфічних норм, адресованих саме системам ШІ.

Стратегічним викликом залишається забезпечення когерентності (англ. – *regulatory coherence*) регіональних регуляторних режимів, що потребує не механічної уніфікації моделей, а розвитку механізмів взаємного визнання та координації між ними задля запобігання фрагментації глобального регуляторного простору та збереження балансу між стимулюванням інновацій і захистом прав людини. У цьому контексті формування універсальної конвенції зі штучного інтелекту залишається стратегічним орієнтиром, досягнення якого залежить від здатності держав дійти консенсусу, забезпечити баланс між правами людини та інтересами технологічного розвитку і національної безпеки, а також створити ефективні механізми моніторингу та забезпечення виконання. Рамкова конвенція Ради Європи, відкрита для приєднання держав з різних регіонів світу, може стати

основою для поступового розширення міжнародного консенсусу до глобального рівня. Водночас геополітична конкуренція між провідними технологічними державами, відмінності у правових традиціях та цінностях, а також економічні інтереси держав та корпорацій у глобальному змаганні за технологічне лідерство створюють значні перешкоди на цьому шляху.

Для України результати дослідження мають практичне значення у кількох вимірах. Порівняльний аналіз регіональних ініціатив підтверджує стратегічну обґрунтованість обраного курсу – засвідченого підписанням Рамкової конвенції Ради Європи 15 травня 2025 року – на її ратифікацію та наближення національного законодавства до вимог Акту про ШІ, тобто актів, що забезпечують захист прав людини у сфері ШІ на рівні юридично обов'язкових норм. Водночас «парадокс глобального регулювання ШІ», за якого регіони та держави з найбільшими потребами у захисті прав людини мають найменші інституційні спроможності для створення ефективних регуляторних рамок, є релевантним і для України, яка в умовах повоєнної відбудови стикатиметься з необхідністю одночасного формування інфраструктури ШІ та створення відповідного правового регулювання. Це вимагатиме не лише ратифікації та імплементації Рамкової конвенції й адаптації законодавства до вимог Акту про ШІ, а й творчого узгодження найкращих практик різних регіональних моделей із реальними інституційними можливостями української держави.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. McCarthy J., Minsky M. L., Rochester N., Shannon C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*. 2006. Vol. 27, no. 4. P. 12–14. DOI: 10.1609/aimag.v27i4.1904. URL: <https://onlinelibrary.wiley.com/doi/epdf/10.1609/aimag.v27i4.1904> (дата звернення: 18.02.2026).
2. Maas M. M. International Law Does Not Compute: Artificial Intelligence and the Development, Displacement or Destruction of the Global Legal Order. *Melbourne Journal of International Law*. 2019. Vol. 20, no. 1. P. 29–106. URL: <https://classic.austlii.edu.au/au/journals/MelbJIL/2019/3.html> (дата звернення: 18.02.2026).
3. Попова Н. О. Міжнародно-правове регулювання штучного інтелекту: сучасний стан та перспективи розвитку. *Юридичний науковий електронний журнал*. 2024. № 4. С. 245–248. DOI: 10.32782/2524-0374/2024-4/105. URL: https://www.lsej.org.ua/4_2024/107.pdf (дата звернення: 18.02.2026).
4. Onishchenko N., Kostenko O., Zhuravlov D. AI Technologies to the Question of the "Policy" of Legal Regulation at the Present Stage. Essential and Instrumental Factors. *International Journal of Innovative Technologies in Social Science*. 2025. No. 1 (45). DOI: 10.31435/ijitss.1(45).2025.3200. URL: <https://rsglobal.pl/index.php/ijitss/article/view/3200> (дата звернення: 18.02.2026).
5. European Parliament. Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) : A8-0005/2017. 2017. URL: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html (дата звернення: 18.02.2026).
6. European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) : P8_TA(2017)0051. 2017. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017IP0051> (дата звернення: 18.02.2026).

7. Guzman A. T., Meyer T. L. International Soft Law. *Journal of Legal Analysis*. 2010. Vol. 2, no. 1. P. 171–225. URL: https://scholarship.law.duke.edu/faculty_scholarship/4207/ (дата звернення: 18.02.2026).
8. OECD AI Policy Observatory : website. URL: <https://oecd.ai/> (дата звернення: 18.02.2026).
9. Recommendation of the Council on Artificial Intelligence : OECD/LEGAL/0449. OECD, 2019 (updated 2024). URL: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf> (дата звернення: 18.02.2026).
10. G20. AI Principles. Ministerial Statement on Trade and Digital Economy. Tsukuba, Japan, 2019. URL: https://g7g20-documents.org/fileadmin/G7G20_documents/2019/G20/Japan/Leaders/2%20Leaders%27%20Annex/G20%20AI%20Principles_2019.pdf (дата звернення: 18.02.2026).
11. United Nations. Secretary-General's Strategy on New Technologies. United Nations, 2018. URL: <https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf> (дата звернення: 18.02.2026).
12. Records of the General Conference, 40th session : Resolutions : 40 C/Resolution 53. Paris : UNESCO, 2019. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000372579> (дата звернення: 18.02.2026).
13. Recommendation on the Ethics of Artificial Intelligence : adopted by the General Conference at its 41st session. Paris : UNESCO, 2021. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000380455> (дата звернення: 18.02.2026).
14. Kaushik A. UNESCO's Recommendation on the Ethics of AI. Montreal AI Ethics Institute, 2021. URL: <https://montrealetics.ai/unescos-recommendation-on-the-ethics-of-ai/> (дата звернення: 18.02.2026).
15. Ігнатуша В. В. Розвиток міжнародно-правового регулювання штучного інтелекту в рамках ООН. Аналітично-порівняльне правознавство. 2025. № 3, ч. 3.

C. 287–296. DOI: 10.24144/2788-6018.2025.03.3.45. URL: <https://journal-app.uzhnu.edu.ua/article/view/335016> (дата звернення: 18.02.2026).

16. Transforming our world: the 2030 Agenda for Sustainable Development : resolution adopted by the General Assembly on 25 September 2015 : A/RES/70/1. URL: https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf (дата звернення: 18.02.2026).

17. The Right to Privacy in the Digital Age : resolution adopted by the General Assembly on 15 December 2022 : A/RES/77/211. URL: <https://documents.un.org/doc/undoc/gen/n22/762/14/pdf/n2276214.pdf> (дата звернення: 18.02.2026).

18. Impact of Rapid Technological Change on the Achievement of the Sustainable Development Goals and Targets : resolution adopted by the General Assembly on 25 July 2023 : A/RES/77/320. URL: <https://documents.un.org/doc/undoc/gen/n23/227/47/pdf/n2322747.pdf> (дата звернення: 18.02.2026).

19. Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development : resolution adopted by the General Assembly on 21 March 2024 : A/RES/78/265. URL: <https://documents.un.org/doc/undoc/gen/n24/087/83/pdf/n2408783.pdf> (дата звернення: 18.02.2026).

20. General Assembly Adopts Landmark Resolution on Artificial Intelligence. UN News. 2024. 21 March. URL: <https://news.un.org/en/story/2024/03/1147831> (дата звернення: 18.02.2026).

21. The Pact for the Future : resolution adopted by the General Assembly on 22 September 2024 : A/RES/79/1. URL: <https://documents.un.org/doc/undoc/ltd/n24/252/91/pdf/n2425291.pdf> (дата звернення: 18.02.2026).

22. Кресін О. В. Резолюції Генеральної Асамблеї ООН: характер і значення у контексті війни РФ проти України : монографія. Київ : Норма права, 2024. – 236 с.

23. Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. Council of Europe, 2020.

URL: <https://rm.coe.int/09000016809e1154> (дата звернення: 18.02.2026).

24. Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225) : adopted 17 May 2024. Council of Europe, 2024. URL: <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=225> (дата звернення: 18.02.2026).

25. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. 2024. L, 2024/1689. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng> (дата звернення: 18.02.2026).

26. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence : COM(2021) 206 final. European Commission, 2021. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206> (дата звернення: 18.02.2026).

27. Artificial Intelligence Act: MEPs adopt landmark law : press release, 13 March 2024 : 20240308IPR19015. European Parliament, 2024. URL:

<https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law> (дата звернення: 18.02.2026).

28. Global Partnership on Artificial Intelligence : founding documents. 2020. URL: <https://gpai.ai/> (дата звернення: 18.02.2026).

29. AI for Good Global Summit. International Telecommunication Union, 2017. URL: <https://aiforgood.itu.int/> (дата звернення: 18.02.2026).

30. Машлій Г. Б., Мосій О., Пельчер М. Дослідження управлінських аспектів використання штучного інтелекту. Галицький економічний вісник. Тернопіль : ТНТУ, 2019. Т. 57, № 2. С. 80—89. URL:

<https://galicianvisnyk.tntu.edu.ua/pdf/57/601.pdf> (дата звернення: 18.02.2026).

31. Scherer M. U. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*. 2016. Vol. 29, no. 2. P. 353–400. DOI: 10.2139/ssrn.2609777. URL:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609777 (дата звернення: 18.02.2026).

32. Корнеєва С. Р. Теоретичні підходи до визначення поняття та правового регулювання штучного інтелекту. *Науковий вісник Ужгородського національного університету. Серія «Право»*. 2021. Вип. 66. С. 50–55. DOI: 10.24144/2307-3322.2021.66.9. URL:

<https://visnyk-juris-uzhnu.com/wp-content/uploads/2021/11/11.pdf> (дата звернення: 18.02.2026).

33. McCarthy J. What is Artificial Intelligence? Stanford University, 2007. URL: <http://www-formal.stanford.edu/jmc/whatisai.pdf> (дата звернення: 18.02.2026).

34. Human Intelligence. *Encyclopedia Britannica*. 2024. URL: <https://www.britannica.com/science/human-intelligence-psychology> (дата звернення: 06.03.2026).

35. Colom R., Karama S., Jung R. E., Haier R. J. Human Intelligence and Brain Networks. *Dialogues in Clinical Neuroscience*. 2010. Vol. 12, no. 4. P. 489–501. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3181994/> (дата звернення: 18.02.2026).

36. Neisser U., Boodoo G., Bouchard T. J. Jr., Boykin A. W., Brody N., Ceci S. J., Halpern D. F., Loehlin J. C., Perloff R., Sternberg R. J., Urbina S. Intelligence: Knowns and Unknowns. *American Psychologist*. 1996. Vol. 51, no. 2. P. 77–101. DOI: 10.1037/0003-066X.51.2.77. URL: https://www.researchgate.net/publication/228356395_Intelligence_Knowns_and_unknowns (дата звернення: 18.02.2026).

37. Sternberg R. J. The Triarchic Theory of Successful Intelligence. *Contemporary Intellectual Assessment: Theories, Tests, and Issues* / eds. D. P. Flanagan, E. M. McDonough. 4th ed. New York : The Guilford Press, 2018. P. 174–194. URL: https://www.researchgate.net/publication/26610484_The_Theory_of_Successful_Intellig

[ence/fulltext/0ffc9e4c0cf255165fc9efe9/The-Theory-of-Successful-Intelligence.pdf](https://www.sciencedirect.com/science/fulltext/0ffc9e4c0cf255165fc9efe9/The-Theory-of-Successful-Intelligence.pdf) (дата звернення: 18.02.2026).

38. Cognition. ScienceDirect Topics. 2011. URL: <https://www.sciencedirect.com/topics/immunology-and-microbiology/cognition> (дата звернення: 18.02.2026).

39. Harvey P. D. Domains of Cognition and Their Assessment. Dialogues in Clinical Neuroscience. 2019. Vol. 21, no. 3. P. 227–237. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6829170/> (дата звернення: 18.02.2026).

40. Fisher G. G., Chaffee D. S., Tetric L. E., Davalos D. B., Potter G. G. Cognitive Functioning, Aging, and Work: A Review and Recommendations for Research and Practice. Journal of Occupational Health Psychology. 2017. Vol. 22, no. 3. P. 314–336. DOI: 10.1037/ocp0000086. URL: <https://psycnet.apa.org/manuscript/2017-14294-001.pdf> (дата звернення: 18.02.2026).

41. Баранов О. А. Визначення терміну «штучний інтелект». Інформація і право. 2023. № 1 (44). С. 32–49. URL: https://ippi.org.ua/sites/default/files/5_28.pdf (дата звернення: 18.02.2026).

42. Hartmann S. The World as a Process: Simulations in the Natural and Social Sciences. Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View / eds. R. Hegselmann et al. Dordrecht : Springer, 2005. P. 77–100. URL: https://www.researchgate.net/publication/222712674_The_World_as_a_Process_Simulations_in_the_Natural_and_Social_Sciences (дата звернення: 19.03.2026).

43. Agüera y Arcas B. Life, Intelligence, and Consciousness: A Functional Perspective. Long Now Ideas. 2025. URL: <https://longnow.org/ideas/life-intelligence-consciousness/> (дата звернення: 18.02.2026).

44. Bianchini F. A New Definition of "Artificial" for Two Artificial Sciences. Foundations of Science. 2023. Vol. 28. P. 401–417. DOI: 10.1007/s10699-021-09799-w. URL: <https://link.springer.com/article/10.1007/s10699-021-09799-w> (дата звернення: 18.02.2026).

45. Sheikh H., Prins C., Schrijvers E. Artificial Intelligence: Definition and Background. Mission AI: The New System Technology. Cham : Springer, 2023. P. 9–22. URL: https://link.springer.com/chapter/10.1007/978-3-031-21448-6_2 (дата звернення: 18.02.2026).

46. Klein E., Patrick S. Envisioning a Global Regime Complex to Govern Artificial Intelligence. Carnegie Endowment for International Peace, 2024. URL: https://carnegie-production-assets.s3.amazonaws.com/static/files/Klein_Patrick_AI_Regime_Complex-1.pdf (дата звернення: 18.02.2026).

47. Ethics Guidelines for Trustworthy AI / European Commission High-Level Expert Group on Artificial Intelligence. European Commission, 2019. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (дата звернення: 18.02.2026).

48. Report on Artificial Intelligence: Questions of Interpretation and Application of International Law (2020/2013(INI)) : A9-0001/2021. European Parliament, 2021. URL: https://www.europarl.europa.eu/doceo/document/A-9-2021-0001_EN.html (дата звернення: 18.02.2026).

49. Про схвалення Концепції розвитку штучного інтелекту в Україні : розпорядження Кабінету Міністрів України від 2 грудня 2020 р. № 1556-р. URL: <https://www.kmu.gov.ua/npas/pro-shvalennya-konceptsiyi-rozvitku-shtuchnogo-intelektu-v-ukrayini-s21220> (дата звернення: 18.02.2026).

50. Goertzel B. Artificial General Intelligence: Concept, State of the Art, and Future Prospects. Journal of Artificial General Intelligence. 2014. Vol. 5, no. 1. P. 1–48. DOI: 10.2478/jagi-2014-0001. URL: <https://reference-global.com/article/10.2478/jagi-2014-0001> (дата звернення: 18.02.2026).

51. Morris M. R., Sohl-Dickstein J., Fiedel N. et al. Levels of AGI for Operationalizing Progress on the Path to AGI. Proceedings of the 41st International Conference on Machine Learning (ICML), Vienna, Austria. PMLR, 2024. Vol. 235.

DOI: 10.48550/arXiv.2311.02462. URL: <https://arxiv.org/pdf/2311.02462> (дата звернення: 18.02.2026).

52. Awad A., Littig L., Geraerds S. Artificial General Intelligence: Demystifying AI's Next Frontier : white paper. Boston Consulting Group, 2024. URL: <https://media-publications.bcg.com/BCG-Artificial-General-Intelligence-Whitepaper.pdf> (дата звернення: 18.02.2026).

53. Gubrud M. A. Nanotechnology and International Security : draft paper for a talk at the Fifth Foresight Conference on Molecular Nanotechnology. Center for Superconductivity Research, 1997. URL: <https://legacy.foresight.org/Conferences/MNT05/Papers/Gubrud/index.html> (дата звернення: 18.02.2026).

54. Bostrom N. Superintelligence: Paths, Dangers, Strategies. Oxford : Oxford University Press, 2014. 328 p. URL: https://dn790007.ca.archive.org/0/items/artificial-intelligence/2014_superintelligence_paths%2C_dangers%2C_strategies-Nick%20Bostrom.pdf (дата звернення: 18.02.2026).

55. Russell S. J., Norvig P. Artificial Intelligence: A Modern Approach. 3rd ed. Upper Saddle River : Prentice Hall, 2010. 1132 p. URL: <https://people.engr.tamu.edu/guni/csce625/slides/AI.pdf> (дата звернення: 18.02.2026).

56. Omohundro S. M. The Basic AI Drives. Artificial General Intelligence 2008 : proceedings of the First AGI Conference / eds. P. Wang, B. Goertzel, S. Franklin. Amsterdam : IOS Press, 2008. P. 483–492. URL: https://selfawareness.com/wp-content/uploads/2008/01/ai_drives_final.pdf (дата звернення: 18.02.2026).

57. OECD Framework for the Classification of AI Systems : OECD Digital Economy Papers, No. 323. Paris : OECD Publishing, 2022. DOI: 10.1787/cb6d9eca-en. URL: https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html (дата звернення: 18.02.2026).

58. Gómez de Ágreda Á. Ethics of Autonomous Weapons Systems and Its Applicability to Any AI Systems. Telecommunications Policy. 2020. Vol. 44, no. 6. Art. 101953. DOI: 10.1016/j.telpol.2020.101953. URL:

<https://www.sciencedirect.com/science/article/pii/S0308596120300458> (дата звернення: 18.02.2026).

59. Views of the ICRC on Autonomous Weapon Systems : paper submitted to the Convention on Certain Conventional Weapons Meeting of Experts on Lethal Autonomous Weapons Systems. International Committee of the Red Cross, 2016. URL: <https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system> (дата звернення: 18.02.2026).

60. Ігнатуша В. В. Міжнародно-правові механізми забезпечення контролю людини над розвитком систем штучного інтелекту. Український часопис міжнародного права. 2025. № 2. С. 67–71. DOI: 10.36952/ujil.2025.2.67-71. URL: <https://jusintergentes.com.ua/images/2025/2/67-71.pdf> (дата звернення: 18.02.2026).

61. Report of the Secretary-General on Current Developments in Science and Technology and Their Potential Impact on International Security : A/79/224. United Nations, 2024. URL: <https://docs.un.org/en/A/79/224> (дата звернення: 18.02.2026).

62. Ігнатуша В. В. Автономні системи з ШІ у збройних конфліктах та правове забезпечення дотримання принципу розрізнення. Науковий вісник Ужгородського національного університету. Серія «Право». 2025. Вип. 90, ч. 5. С. 151–161. DOI: 10.24144/2307-3322.2025.90.5.18. URL: <https://visnyk-pravo.uzhnu.edu.ua/article/view/340844> (дата звернення: 18.02.2026).

63. White Paper on Artificial Intelligence – A European Approach to Excellence and Trust : COM(2020) 65 final. European Commission, 2020. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0065> (дата звернення: 18.02.2026).

64. Brożek B., Furman M., Jakubiec M., Kucharzyk B. The Black Box Problem Revisited. Real and Imaginary Challenges for Automated Legal Decision Making. Artificial Intelligence and Law. 2023. Vol. 32. P. 1–14. DOI: 10.1007/s10506-023-09356-9. URL: https://www.researchgate.net/publication/369791232_The_black_box_problem_revisited

Real and imaginary challenges for automated legal decision making (дата звернення: 18.02.2026).

65. Bathaee Y. The Artificial Intelligence Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology*. 2018. Vol. 31, no. 2. P. 889–938. URL: <https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf> (дата звернення: 18.02.2026).

66. Pavlidis G. Unlocking the Black Box: Analysing the EU Artificial Intelligence Act's Framework for Explainability in AI. *Law, Innovation and Technology*. 2024. Vol. 16, no. 1. P. 293–308. DOI: 10.1080/17579961.2024.2313795. URL: <https://arxiv.org/pdf/2502.14868> (дата звернення: 18.02.2026).

67. Белов Д. М., Белова М. В., Рушак І. В. Штучний інтелект у досудовому розслідуванні кримінальних справ: окремі питання міжнародної практики. *Аналітично-порівняльне правознавство*. 2025. № 1. С. 818–824. DOI: 10.24144/2788-6018.2025.01.136. URL: <https://journal-app.uzhnu.edu.ua/article/view/324040> (дата звернення: 18.02.2026).

68. Камінська Н., Камінська А. Новелізація правового регулювання штучного інтелекту: міжнародний і наднаціональний виміри. *Наше право*. 2025. № 2. С. 234–242. URL: https://nashe-pravo.unesco-socio.in.ua/wp-content/uploads/archive/NP-2025-2/NP_2025_2_234.pdf (дата звернення: 18.02.2026).

69. Фалалєєва Л. Г. Взаємодія парадигм міжнародного права прав людини і міжнародного гуманітарного права крізь призму обов'язків окупаційної держави. *Правова держава : щорічник наукових праць Інституту держави і права імені В. М. Корецького НАН України*. Київ, 2023. Вип. 34. С. 582–598. DOI: 10.33663/1563-3349-2023-34-582-598. URL: <http://pravova-derzhava.org.ua/ua/zmist-nomeriv/pravova-derjava.-vipusk-34-2023.html> (дата звернення: 18.02.2026).

70. Draft Articles on Responsibility of States for Internationally Wrongful Acts, with Commentaries : adopted by the International Law Commission at its fifty-third session, 2001. *Yearbook of the International Law Commission*, 2001. Vol. II, part 2.

URL: https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf (дата звернення: 18.02.2026).

71. The Antarctic Treaty : signed at Washington on 1 December 1959 : entered into force on 23 June 1961. United Nations Treaty Series. Vol. 402. No. 5778. URL: <https://www.ats.aq/e/antarctic treaty.html> (дата звернення: 18.02.2026).

72. Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies : adopted by the General Assembly in resolution 2222 (XXI) of 19 December 1966 : opened for signature on 27 January 1967 : entered into force on 10 October 1967. United Nations Treaty Series. Vol. 610. No. 8843. URL: <https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html> (дата звернення: 18.02.2026).

73. United Nations Convention on the Law of the Sea : signed at Montego Bay on 10 December 1982. United Nations Treaty Series. Vol. 1833. No. 31363. P. 397. URL: https://www.un.org/depts/los/convention_agreements/texts/unclos/unclos_e.pdf (дата звернення: 18.02.2026).

74. Cartagena Protocol on Biosafety to the Convention on Biological Diversity : Montreal, 29 January 2000. URL: <https://bch.cbd.int/protocol/text/> (дата звернення: 18.02.2026).

75. Montreal Protocol on Substances that Deplete the Ozone Layer : Montreal, 16 September 1987. URL: <https://ozone.unep.org/treaties/montreal-protocol> (дата звернення: 18.02.2026).

76. Rio Declaration on Environment and Development : UN Conference on Environment and Development, Rio de Janeiro, 3–14 June 1992 : A/CONF.151/26 (Vol. I). URL: https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf (дата звернення: 18.02.2026).

77. Matthias A. The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata. Ethics and Information Technology. 2004. Vol. 6, no. 3. P. 175–

183. DOI: 10.1007/s10676-004-3422-1. URL:

<https://link.springer.com/article/10.1007/s10676-004-3422-1> (дата звернення: 18.02.2026).

78. Collingridge D. The Social Control of Technology. London : Frances Pinter, 1980. 200 p. URL: <https://archive.org/details/socialcontroloft0000coll/mode/2up> (дата звернення: 18.02.2026).

79. Downes L., Mui C. Unleashing the Killer App: Digital Strategies for Market Dominance. Boston : Harvard Business School Press, 2000. 256 p. URL: https://vintageapple.org/apple_jobs/pdf/Unleashing_the_Killer_App_Downes_2000.pdf (дата звернення: 18.02.2026).

80. Fjeld J., Achten N., Hilligoss H., Nagy A., Srikumar M. Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI : Berkman Klein Center Research Publication No. 2020-1. Harvard University, 2020. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518482 (дата звернення: 18.02.2026).

81. Chinen M. The International Governance of Artificial Intelligence. Cheltenham : Edward Elgar Publishing, 2023. 384 p. ISBN 978-1-80037-922-9.

82. Outcome Document of the High-Level Meeting of the General Assembly on the Overall Review of the Implementation of the Outcomes of the World Summit on the Information Society : resolution adopted by the General Assembly on 16 December 2015 : A/RES/70/125. URL: <https://digitallibrary.un.org/record/819076?ln=en&v=pdf> (дата звернення: 18.02.2026).

83. Information and Communications Technologies for Sustainable Development : resolution adopted by the General Assembly on 19 December 2023 : A/RES/78/132. URL: <https://digitallibrary.un.org/record/4031524?ln=en&v=pdf> (дата звернення: 18.02.2026).

84. Science, Technology and Innovation for Sustainable Development : resolution adopted by the General Assembly on 19 December 2023 : A/RES/78/160. URL: <https://digitallibrary.un.org/record/4031991?ln=en&v=pdf> (дата звернення: 18.02.2026).

85. Lethal Autonomous Weapons Systems : resolution adopted by the General Assembly on 22 December 2023 : A/RES/78/241. URL: <https://digitallibrary.un.org/record/4033027?ln=en&v=pdf> (дата звернення: 18.02.2026).

86. Enhancing International Cooperation on Capacity-Building of Artificial Intelligence : resolution adopted by the General Assembly on 1 July 2024 : A/RES/78/311. URL: <https://digitallibrary.un.org/record/4054005?ln=en&v=pdf> (дата звернення: 18.02.2026).

87. Lethal Autonomous Weapons Systems : resolution adopted by the General Assembly on 2 December 2024 : A/RES/79/62. URL: <https://digitallibrary.un.org/record/4071100?ln=en&v=pdf> (дата звернення: 18.02.2026).

88. Information and Communications Technologies for Sustainable Development : resolution adopted by the General Assembly on 19 December 2024 : A/RES/79/194. URL: <https://digitallibrary.un.org/record/4072126?ln=en&v=pdf> (дата звернення: 18.02.2026).

89. Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security : resolution adopted by the General Assembly on 24 December 2024 : A/RES/79/239. URL: <https://digitallibrary.un.org/record/4071348?ln=en&v=pdf> (дата звернення: 18.02.2026).

90. Role of Artificial Intelligence in Creating New Opportunities for Sustainable Development in Central Asia : resolution adopted by the General Assembly on 25 July 2025 : A/RES/79/322. URL: <https://digitallibrary.un.org/record/4086743?ln=en&v=pdf> (дата звернення: 18.02.2026).

91. Terms of Reference and Modalities for the Establishment and Functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance : resolution adopted by the General Assembly on 26 August 2025 : A/RES/79/325. URL: <https://digitallibrary.un.org/record/4087699> (дата звернення: 18.02.2026).

92. Impact of Rapid Technological Change on the Achievement of the Sustainable Development Goals and Targets : resolution adopted by the General Assembly on 5

September 2025 : A/RES/79/334. URL: <https://digitallibrary.un.org/record/4088140> (дата звернення: 18.02.2026).

93. Possible Risks of the Integration of Artificial Intelligence into Command, Control and Communications Systems of Nuclear Weapons : resolution adopted by the General Assembly on 1 December 2025 : A/RES/80/23. URL: <https://digitallibrary.un.org/record/4095461> (дата звернення: 18.02.2026).

94. Lethal Autonomous Weapons Systems : resolution adopted by the General Assembly on 1 December 2025 : A/RES/80/57. URL: <https://digitallibrary.un.org/record/4095989> (дата звернення: 18.02.2026).

95. Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security : resolution adopted by the General Assembly on 1 December 2025 : A/RES/80/58. URL: <https://digitallibrary.un.org/record/4095956> (дата звернення: 18.02.2026)

96. Information in the Service of Humanity : resolution adopted by the General Assembly on 5 December 2025 : A/RES/80/84 A–B. URL: <https://digitallibrary.un.org/record/4096871> (дата звернення: 18.02.2026)

97. Information and Communications Technologies for Sustainable Development : resolution adopted by the General Assembly on 15 December 2025 : A/RES/80/118. URL: <https://digitallibrary.un.org/record/4097019> (дата звернення: 18.02.2026).

98. Science, Technology and Innovation for Sustainable Development : resolution adopted by the General Assembly on 15 December 2025 : A/RES/80/147. URL: <https://digitallibrary.un.org/record/4097730> (дата звернення: 18.02.2026).

99. Outcome Document of the High-Level Meeting of the General Assembly on the Overall Review of the Implementation of the Outcomes of the World Summit on the Information Society : resolution adopted by the General Assembly on 17 December 2025 : A/RES/80/173. URL: <https://digitallibrary.un.org/record/4097834> (дата звернення: 18.02.2026).

100. Natorski M. Multilateralism in the Global Governance of Artificial Intelligence. 2025. DOI: 10.48550/arXiv.2508.15397. URL:

https://www.researchgate.net/publication/394831494_Multilateralism_in_the_Global_Governance_of_Artificial_Intelligence (дата звернення: 18.02.2026).

101. Vercelli A. UN General Assembly AI Resolutions: 78/265 and 78/311. Proceedings of the 3rd Workshop on Bias, Ethical AI, Explainability and the Role of Logic and Logic Programming (BEWARE24), Bolzano, Italy, 25–28 November 2024. CEUR-WS, 2024. Vol. 3881. URL: <https://ceur-ws.org/Vol-3881/paper11.pdf> (дата звернення: 18.02.2026).

102. Hao K. AI's Growing Carbon Footprint. Columbia Climate School, 2023. URL: <https://news.climate.columbia.edu/2023/06/09/ais-growing-carbon-footprint/> (дата звернення: 18.02.2026).

103. The AI Index 2022 Annual Report. Stanford University, Human-Centered Artificial Intelligence, 2022. URL: https://hai.stanford.edu/assets/files/2022-ai-index-report_master.pdf (дата звернення: 18.02.2026).

104. Dwyer A. A Tale of Two Shortages: Reconciling Demand for Water and Microchips in Arizona. American Bar Association, Section of Environment, Energy, and Resources, 2023. URL: https://www.americanbar.org/groups/environment_energy_resources/resources/newsletters/water/reconciling-demand-water-microchips-arizona/ (дата звернення: 18.02.2026).

105. How Can We Reduce Environmental Impact in Chip Manufacturing? Imec, 2024. URL: <https://www.imec-int.com/en/articles/how-can-we-reduce-environmental-impact-chip-manufacturing> (дата звернення: 18.02.2026).

106. Readiness Assessment Methodology for the Recommendation on the Ethics of Artificial Intelligence. Paris : UNESCO, 2022. URL: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/readiness-assessment> (дата звернення: 18.02.2026).

107. Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence. Paris : UNESCO, 2023. URL:

<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics> (дата звернення: 18.02.2026).

108. Ігнатуша В. В. Ризики використання штучного інтелекту в оборонній сфері та способи їх подолання в рамках ЄС та НАТО. Інтеграція України до Європейського Союзу: виклики в умовах воєнного стану. Пам'яті професора Павла Заворотька та професора Володимира Денисова : зб. матеріалів і тез доп. міжнар. наук.-практ. конф. (м. Київ, 4–6 листопада 2024 р.) / за наук. ред. проф. Л. Г. Фалалєєвої. Київ : Вид-во «Людмила», 2025. С. 299–303. URL:

<https://idpnan.org/integration-into-european-union> (дата звернення: 18.02.2026).

109. Resolution 101 – Standardization Activities of the ITU Telecommunication Standardization Sector on Artificial Intelligence Technologies in Support of Telecommunications/Information and Communication Technologies : World Telecommunication Standardization Assembly, New Delhi, 15–24 October 2024. Geneva : ITU, 2024. URL: https://www.itu.int/dms_pub/itu-t/opb/res/T-RES-T.101-2024-PDF-E.pdf (дата звернення: 18.02.2026).

110. Standardization Roadmap on AI for Natural Disaster Management: Trends and Gaps in Standardization : Technical Report FG-AI4NDM / ITU/WMO/UNEP Focus Group on AI for Natural Disaster Management. Geneva : ITU, 2023. URL:

<https://www.itu.int/en/ITU-T/focusgroups/ai4ndm/Pages/default.aspx> (дата звернення: 18.02.2026).

111. Birkstedt T., Minkkinen M., Tandon A., Mäntymäki M. AI Governance: Themes, Knowledge Gaps and Future Agendas. Internet Research. 2023. Vol. 33, no. 7. P. 133–167. DOI: 10.1108/INTR-01-2022-0042. URL:

<https://www.emerald.com/intr/article/33/7/133/178343/AI-governance-themes-knowledge-gaps-and-future> (дата звернення: 18.02.2026).

112. Resolution 214 – Artificial Intelligence Technologies and Telecommunications/Information and Communication Technologies : Plenipotentiary Conference, Bucharest, 2022. Geneva : ITU, 2022. URL:

<https://www.itu.int/en/council/Documents/basic-texts-2023/RES-214-E.pdf> (дата звернення: 18.02.2026).

113. Ethics and Governance of Artificial Intelligence for Health : WHO Guidance. Geneva : World Health Organization, 2021. 165 p. ISBN 978-92-4-002920-0. URL: <https://www.who.int/publications/i/item/9789240029200> (дата звернення: 18.02.2026).

114. Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multi-Modal Models. Geneva : World Health Organization, 2024. ISBN 978-92-4-008475-9. URL: <https://www.who.int/publications/i/item/9789240084759> (дата звернення: 18.02.2026).

115. Regulatory Considerations on Artificial Intelligence for Health. Geneva : World Health Organization, 2023. ISBN 978-92-4-007887-1. URL: <https://www.who.int/publications/i/item/9789240078871> (дата звернення: 18.02.2026).

116. Global Strategy on Digital Health 2020–2025. Geneva : World Health Organization, 2021. ISBN 978-92-4-002092-4. URL: <https://www.who.int/publications/i/item/9789240020924> (дата звернення: 18.02.2026).

117. Generating Evidence for Artificial Intelligence-Based Medical Devices: A Framework for Training, Validation and Evaluation. Geneva : World Health Organization, 2021. ISBN 978-92-4-005987-0. URL: <https://www.who.int/publications/i/item/9789240059870> (дата звернення: 18.02.2026).

118. WIPO Technology Trends 2019: Artificial Intelligence. Geneva : World Intellectual Property Organization, 2019. ISBN 978-92-805-3004-8. URL: <https://www.wipo.int/publications/en/details.jsp?id=4386> (дата звернення: 18.02.2026).

119. Draft Issues Paper on Intellectual Property Policy and Artificial Intelligence : WIPO Conversation on IP and AI, Second Session : WIPO/IP/AI/2/GE/20/1. Geneva : World Intellectual Property Organization, 2020. URL: https://www.wipo.int/about-ip/en/artificial_intelligence/conversation.html (дата звернення: 18.02.2026).

120. Generative AI: Navigating Intellectual Property. Geneva : World Intellectual Property Organization, 2024. URL:

https://www.wipo.int/about-ip/en/artificial_intelligence/generative_ai.html (дата звернення: 18.02.2026).

121. WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI) : session reports. Geneva : World Intellectual Property Organization, 2019. URL: https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=459091 (дата звернення: 18.02.2026).

122. AI and IP Strategy Clearing House : database of national AI and IP strategies. Geneva : World Intellectual Property Organization, 2020. URL: https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html (дата звернення: 18.02.2026).

123. Gmyrek P., Berg J., Bescond D. Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality : ILO Working Paper 96. Geneva : International Labour Organization, 2023. DOI: 10.54394/FHEM8239. URL: <https://researchrepository.ilo.org/esploro/outputs/encyclopediaEntry/995326516102676> (дата звернення: 18.02.2026).

124. Gmyrek P., Berg J., Kamiński K., Konopczyński F., Ładna A., Nafradi B., Rosłaniec K., Troszyński M. Generative AI and Jobs: A Refined Global Index of Occupational Exposure : ILO Working Paper 140. Geneva : International Labour Organization, 2025. URL: <https://www.ilo.org/publications/generative-ai-and-jobs-refined-global-index-occupational-exposure> (дата звернення: 18.02.2026).

125. Work Transformed: The Promise and Peril of Artificial Intelligence : analytical brief. Geneva : International Labour Organization, 2025. URL: <https://www.ilo.org/publications/work-transformed-promise-and-peril-artificial-intelligence> (дата звернення: 18.02.2026).

126. Gmyrek P., Winkler H., Garganta S. Buffer or Bottleneck? Employment Exposure to Generative AI and the Digital Divide in Latin America : ILO Working Paper 121. Geneva : International Labour Organization, 2024. DOI: 10.54394/TFZY7681. URL: <https://researchrepository.ilo.org/esploro/outputs/encyclopediaEntry/995379093202676> (дата звернення: 18.02.2026).

127. Observatory on AI and Work in the Digital Economy. International Labour Organization, 2024. URL: <https://www.ilo.org/digital-economy/areas-of-work/ai-and-work> (дата звернення: 18.02.2026).

128. United Nations System White Paper on AI Governance: An Analysis of the UN System's Institutional Models, Functions, and Existing International Normative Frameworks Applicable to AI Governance : CEB/2024/1/Add.1 / Chief Executives Board for Coordination, Inter-Agency Working Group on Artificial Intelligence. New York : United Nations, 2024. URL: <https://unsceb.org/white-paper-ai-governance> (дата звернення: 18.02.2026).

129. UNDP Digital Strategy 2022–2025. New York : United Nations Development Programme, 2022. URL: <https://www.undp.org/publications/undp-digital-strategy-2022-2025> (дата звернення: 18.02.2026).

130. Pujari S., Singh R., Soon G. C. et al. Artificial Intelligence in Traditional Medicine: Policy and Governance Strategies. Bulletin of the World Health Organization. 2025. Vol. 103, no. 11. P. 738–740. DOI: 10.2471/BLT.24.292888. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12578530/> (дата звернення: 18.02.2026).

131. Pujari S., Reis A., Zhao Y. et al. Artificial Intelligence for Global Health: Cautious Optimism with Safeguards. Bulletin of the World Health Organization. 2023. Vol. 101, no. 6. P. 364–364A. DOI: 10.2471/BLT.23.290215. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10225938/> (дата звернення: 18.02.2026).

132. Schiff D. S. Strategies for Harmonizing Fragmented AI Ethics Frameworks, Standards, and Regulations. Handbook of Human-Centered Artificial Intelligence / ed. W. Xu. Singapore : Springer, 2025. DOI: 10.1007/978-981-97-8440-0_82-1. URL: https://link.springer.com/rwe/10.1007/978-981-97-8440-0_82-1 (дата звернення: 18.02.2026).

133. UNESCO AI Readiness Assessment Report: Anchoring Ethics in AI Governance in the Philippines. Paris : UNESCO, 2025. URL: <https://www.unesco.org/en/articles/unesco-ai-readiness-assessment-report-anchoring-ethics-ai-governance-philippines> (дата звернення: 18.02.2026).

134. Corrêa N. K., Galvão C., Santos J. W. et al. Worldwide AI Ethics: A Review of 200 Guidelines and Recommendations for AI Governance. *Patterns*. 2023. Vol. 4, no. 10. Art. 100857. DOI: 10.1016/j.patter.2023.100857. URL:

[https://www.cell.com/patterns/fulltext/S2666-3899\(23\)00241-6](https://www.cell.com/patterns/fulltext/S2666-3899(23)00241-6) (дата звернення: 18.02.2026).

135. Krijger J., Thuis T., de Ruyter M., Ligthart E., Broekman I. The AI Ethics Maturity Model: A Holistic Approach to Advancing Ethical Data Science in Organizations. *AI and Ethics*. 2023. Vol. 3. P. 355–367. DOI: 10.1007/s43681-022-00228-7. URL: <https://link.springer.com/article/10.1007/s43681-022-00228-7> (дата звернення: 18.02.2026).

136. Scoping the OECD AI Principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO) : OECD Digital Economy Papers, No. 291. Paris : OECD Publishing, 2019. DOI: 10.1787/d62f618a-en. URL:

https://www.oecd.org/en/publications/scoping-the-oecd-ai-principles_d62f618a-en.html (дата звернення: 18.02.2026).

137. Tools for Trustworthy AI: A Framework to Compare Implementation Tools for Trustworthy AI Systems : OECD Digital Economy Papers, No. 312. Paris : OECD Publishing, 2021. DOI: 10.1787/008232ec-en. URL:

https://www.oecd.org/en/publications/tools-for-trustworthy-ai_008232ec-en.html (дата звернення: 18.02.2026).

138. The State of Implementation of the OECD AI Principles Four Years On : OECD Artificial Intelligence Papers, No. 3. Paris : OECD Publishing, 2023. DOI: 10.1787/835641c9-en. URL: https://www.oecd.org/en/publications/the-state-of-implementation-of-the-oecd-ai-principles-four-years-on_835641c9-en.html (дата звернення: 18.02.2026).

139. Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint : OECD Digital Economy Papers, No. 341. Paris : OECD Publishing, 2022. DOI: 10.1787/7babf571-en. URL:

https://www.oecd.org/en/publications/measuring-the-environmental-impacts-of-artificial-intelligence-compute-and-applications_7babf571-en.html (дата звернення: 18.02.2026).

140. Regulatory Sandboxes in Artificial Intelligence : OECD Digital Economy Papers, No. 356. Paris : OECD Publishing, 2023. DOI: 10.1787/8f80a0e6-en. URL: https://www.oecd.org/en/publications/regulatory-sandboxes-in-artificial-intelligence_8f80a0e6-en.html (дата звернення: 18.02.2026).

141. AI, Data Governance and Privacy: Synergies and Areas of International Cooperation : OECD Artificial Intelligence Papers, No. 22. Paris : OECD Publishing, 2024. DOI: 10.1787/2476b1a4-en. URL: https://www.oecd.org/en/publications/ai-data-governance-and-privacy_2476b1a4-en.html (дата звернення: 18.02.2026).

142. Schmitt L. Mapping Global AI Governance: A Nascent Regime in a Fragmented Landscape. *AI and Ethics*. 2021. Vol. 1. P. 303–314. DOI: 10.1007/s43681-021-00083-y. URL: <https://link.springer.com/article/10.1007/s43681-021-00083-y> (дата звернення: 18.02.2026).

143. Roberts H., Hine E., Taddeo M., Floridi L. Understanding the Policy Ecosystem for Governing Artificial Intelligence in International Contexts. *International Affairs*. 2024. Vol. 100, no. 3. P. 1275–1286. DOI: 10.1093/ia/iiae073. URL: <https://academic.oup.com/ia/article/100/3/1275/7641064> (дата звернення: 18.02.2026).

144. Feldstein S. Evaluating Europe's Push to Enact AI Regulations: How Will This Influence Global Norms? *Democratization*. 2024. Vol. 31, no. 5. P. 1049–1066. DOI: 10.1080/13510347.2023.2196068. URL: <https://www.tandfonline.com/doi/full/10.1080/13510347.2023.2196068> (дата звернення: 18.02.2026).

145. Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems. G7, 2023. URL: <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system> (дата звернення: 18.02.2026).

146. Tallberg J., Erman E., Furendal M., Geith J., Klamberg M., Lundgren M. The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative

Research. International Studies Review. 2023. Vol. 25, no. 3. Art. viad040. DOI: 10.1093/isr/viad040. URL: <https://academic.oup.com/isr/article/25/3/viad040/7259354> (дата звернення: 18.02.2026).

147. Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems. G7, 2023. URL: <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems> (дата звернення: 18.02.2026).

148. Cihon P., Maas M. M., Kemp L. Fragmentation and the Future: Investigating Architectures for International AI Governance. Global Policy. 2020. Vol. 11, no. 5. P. 545–556. DOI: 10.1111/1758-5899.12890. URL: <https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.12890> (дата звернення: 18.02.2026).

149. G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI. Paris : OECD Publishing, 2023. DOI: 10.1787/bf3c0c60-en. URL: https://www.oecd.org/en/publications/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en.html (дата звернення: 18.02.2026).

150. Roberts H., Taddeo M., Floridi L. A Framework for Evaluating Global AI Governance Initiatives : SSRN Working Paper. 2024. DOI: 10.2139/ssrn.5367969. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5367969 (дата звернення: 18.02.2026).

151. Jelinek T., Wallach W., Kerimi D. Policy Brief: The Creation of a G20 Coordinating Committee for the Governance of Artificial Intelligence. AI and Ethics. 2020. Vol. 1. P. 141–150. DOI: 10.1007/s43681-020-00019-y. URL: <https://link.springer.com/article/10.1007/s43681-020-00019-y> (дата звернення: 18.02.2026).

152. Communication from the Commission – Artificial Intelligence for Europe : COM(2018) 237 final, 25 April 2018. European Commission. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0237> (дата звернення: 18.02.2026).

153. European Council Meeting (19 October 2017) – Conclusions : EUCO 14/17. Brussels : European Council, 2017. URL:

<https://www.consilium.europa.eu/media/21620/19-euco-final-conclusions-en.pdf> (дата звернення: 18.02.2026).

154. European Council Conclusions, 28 June 2018 : EUCO 9/18. Brussels : European Council, 2018. URL: <https://www.consilium.europa.eu/media/35936/28-euco-final-conclusions-en.pdf> (дата звернення: 18.02.2026).

155. Communication from the Commission – Coordinated Plan on Artificial Intelligence : COM(2018) 795 final, 7 December 2018. European Commission. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0795> (дата звернення: 18.02.2026).

156. Council Conclusions on the Coordinated Plan on Artificial Intelligence – Adoption : 6177/19. Brussels : Council of the European Union, 2019. URL: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/en/pdf> (дата звернення: 18.02.2026).

157. Communication from the Commission – Building Trust in Human-Centric Artificial Intelligence : COM(2019) 168 final, 8 April 2019. European Commission. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0168> (дата звернення: 18.02.2026).

158. Policy and Investment Recommendations for Trustworthy Artificial Intelligence / European Commission High-Level Expert Group on Artificial Intelligence. European Commission, 2019. URL: <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence> (дата звернення: 18.02.2026).

159. Communication from the Commission – Fostering a European Approach to Artificial Intelligence : COM(2021) 205 final, 21 April 2021. European Commission. URL: <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review> (дата звернення: 18.02.2026).

160. Commission Decision of 24 January 2024 Establishing the European Artificial Intelligence Office : C/2024/1459. Official Journal of the European Union. 2024. C, 14 February 2024. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:C_202401459 (дата звернення: 18.02.2026).

161. AI Pact. Digital Strategy. European Commission, 2024. URL: <https://digital-strategy.ec.europa.eu/en/policies/ai-pact> (дата звернення: 18.02.2026).

162. Standardisation of the AI Act. Digital Strategy. European Commission, 2024. URL: <https://digital-strategy.ec.europa.eu/en/policies/standardisation-ai-act> (дата звернення: 18.02.2026).

163. Draft Implementing Decision Amending Implementing Decision C(2023)3215 on a Standardisation Request to the European Standardisation Organisations in Support of Regulation (EU) 2024/1689 : 14 January 2025. European Commission. URL: <https://ec.europa.eu/transparency/documents-register/> (дата звернення: 18.02.2026).

164. EU Launches InvestAI Initiative to Mobilise €200 Billion of Investment in Artificial Intelligence : press release, 11 February 2025. European Commission. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_25_380 (дата звернення: 18.02.2026).

165. Communication from the Commission – AI Continent Action Plan : COM(2025) 165 final, 9 April 2025. European Commission. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2025:165:FIN> (дата звернення: 18.02.2026).

166. General-Purpose AI Code of Practice – Transparency Chapter. European AI Office, 2025. URL: <https://ec.europa.eu/newsroom/dae/redirection/document/118120> (дата звернення: 18.02.2026).

167. General-Purpose AI Code of Practice – Copyright Chapter. European AI Office, 2025. URL: <https://ec.europa.eu/newsroom/dae/redirection/document/118115> (дата звернення: 18.02.2026).

168. General-Purpose AI Code of Practice – Safety and Security Chapter. European AI Office, 2025. URL: <https://ec.europa.eu/newsroom/dae/redirection/document/118119> (дата звернення: 18.02.2026).

169. Model Documentation Form : annex to the General-Purpose AI Code of Practice – Transparency Chapter. European AI Office, 2025. URL: <https://digital-strategy.ec.europa.eu/en/library/general-purpose-ai-code-practice> (дата звернення: 18.02.2026).

170. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union. 2016. L 119, 4 May 2016. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (дата звернення: 18.02.2026).

171. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC. Official Journal of the European Union. 2019. L 130, 17 May 2019. URL: <https://eur-lex.europa.eu/eli/dir/2019/790/oj> (дата звернення: 18.02.2026).

172. Ігнатуша В. В. Правове регулювання принципів прозорості та підзвітності систем ШІ в ЄС: сучасний стан та перспективи розвитку. Юридичний науковий електронний журнал. 2026. № 1. С. 265–272. DOI: 10.32782/2524-0374/2026-1/58. URL: https://www.lsej.org.ua/1_2026/60.pdf (дата звернення: 18.02.2026).

173. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act). Official Journal of the European Union. 2022. L 152, 3 June 2022. URL: <https://eur-lex.europa.eu/eli/reg/2022/868/oj> (дата звернення: 18.02.2026).

174. Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act). Official Journal of the European Union. 2022. L 265, 12 October 2022. URL:

<https://eur-lex.europa.eu/eli/reg/2022/1925/oj> (дата звернення: 18.02.2026).

175. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act). Official Journal of the European Union. 2022. L 277, 27 October 2022. URL: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj> (дата звернення: 18.02.2026).

176. Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on Measures for a High Common Level of Cybersecurity Across the Union, Amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and Repealing Directive (EU) 2016/1148 (NIS 2 Directive). Official Journal of the European Union. 2022. L 333, 27 December 2022. URL:

<https://eur-lex.europa.eu/eli/dir/2022/2555/oj> (дата звернення: 18.02.2026).

177. Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on Digital Operational Resilience for the Financial Sector and Amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 (Digital Operational Resilience Act). Official Journal of the European Union. 2022. L 333, 27 December 2022. URL: [https://eur-](https://eur-lex.europa.eu/eli/reg/2022/2554/oj)

[lex.europa.eu/eli/reg/2022/2554/oj](https://eur-lex.europa.eu/eli/reg/2022/2554/oj) (дата звернення: 18.02.2026).

178. Decision (EU) 2022/2481 of the European Parliament and of the Council of 14 December 2022 Establishing the Digital Decade Policy Programme 2030. Official Journal of the European Union. 2022. L 323, 19 December 2022. URL: [https://eur-](https://eur-lex.europa.eu/eli/dec/2022/2481/oj)

[lex.europa.eu/eli/dec/2022/2481/oj](https://eur-lex.europa.eu/eli/dec/2022/2481/oj) (дата звернення: 18.02.2026).

179. European Declaration on Digital Rights and Principles for the Digital Decade. Official Journal of the European Union. 2023. C 23, 23 January 2023. URL:

[lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023C0123\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023C0123(01)) (дата звернення: 18.02.2026).

180. Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on General Product Safety, Amending Regulation (EU) No 1025/2012 and Directive (EU) 2020/1828, and Repealing Directive 2001/95/EC (General Product Safety Regulation). Official Journal of the European Union. 2023. L 135, 23 May 2023. URL: <https://eur-lex.europa.eu/eli/reg/2023/988/oj> (дата звернення: 18.02.2026).

181. Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on Machinery and Repealing Directive 2006/42/EC (Machinery Regulation). Official Journal of the European Union. 2023. L 165, 29 June 2023. URL: <https://eur-lex.europa.eu/eli/reg/2023/1230/oj> (дата звернення: 18.02.2026).

182. Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on Harmonised Rules on Fair Access to and Use of Data and Amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act). Official Journal of the European Union. 2023. L, 22 December 2023. URL: <https://eur-lex.europa.eu/eli/reg/2023/2854/oj> (дата звернення: 18.02.2026).

183. Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 Amending Regulation (EU) No 910/2014 as Regards Establishing the European Digital Identity Framework. Official Journal of the European Union. 2024. L, 30 April 2024. URL: <https://eur-lex.europa.eu/eli/reg/2024/1183/oj> (дата звернення: 18.02.2026).

184. Directive (EU) 2024/2831 of the European Parliament and of the Council of 23 October 2024 on Improving Working Conditions in Platform Work. Official Journal of the European Union. 2024. L, 11 November 2024. URL: <https://eur-lex.europa.eu/eli/dir/2024/2831/oj> (дата звернення: 18.02.2026).

185. Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on Liability for Defective Products and Repealing Council Directive 85/374/EEC (Product Liability Directive). Official Journal of the European Union. 2024.

L, 18 November 2024. URL: <https://eur-lex.europa.eu/eli/dir/2024/2853/oj> (дата звернення: 18.02.2026).

186. Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on Horizontal Cybersecurity Requirements for Products with Digital Elements and Amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act). Official Journal of the European Union. 2024. L, 20 November 2024. URL: <https://eur-lex.europa.eu/eli/reg/2024/2847/oj> (дата звернення: 18.02.2026).

187. Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and Amending Directive 2011/24/EU and Regulation (EU) 2024/2847 (European Health Data Space Regulation). Official Journal of the European Union. 2025. L, 5 March 2025. URL: <https://eur-lex.europa.eu/eli/reg/2025/327/oj> (дата звернення: 18.02.2026).

188. Laux J., Wachter S., Mittelstadt B. Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and Acceptability of Risk. Regulation & Governance. 2024. Vol. 18, no. 1. P. 3–32. DOI: 10.1111/rego.12512. URL: <https://onlinelibrary.wiley.com/doi/10.1111/rego.12512> (дата звернення: 18.02.2026).

189. Schuett J. Risk Management in the Artificial Intelligence Act. European Journal of Risk Regulation. 2024. Vol. 15, no. 2. P. 367–385. DOI: 10.1017/err.2023.1. URL: <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/risk-management-in-the-artificial-intelligence-act/2E4D5707E65EFB3251A76E288BA74068> (дата звернення: 18.02.2026).

190. Novelli C., Casolari F., Rotolo A., Taddeo M., Floridi L. AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act. Digital Society. 2024. Vol. 3. Art. 13. DOI: 10.1007/s44206-024-00095-1. URL: <https://link.springer.com/article/10.1007/s44206-024-00095-1> (дата звернення: 18.02.2026).

191. Lazcoz G., de Hert P. Humans in the GDPR and AIA Governance of Automated and Algorithmic Systems. Essential Pre-Requisites Against Abdicating Responsibilities. *Computer Law & Security Review*. 2023. Vol. 50. Art. 105833. DOI: 10.1016/j.clsr.2023.105833. URL: <https://www.sciencedirect.com/science/article/pii/S0267364923000432> (дата звернення: 18.02.2026).

192. Wachter S. Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond. *Yale Journal of Law & Technology*. 2024. Vol. 26, no. 3. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4924553 (дата звернення: 18.02.2026).

193. Almada M., Radu A. The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy. *German Law Journal*. 2024. Vol. 25. P. 646–663. DOI: 10.1017/glj.2023.108. URL: <https://www.cambridge.org/core/journals/german-law-journal/article/brussels-sideeffect-how-the-ai-act-can-reduce-the-global-reach-of-eu-policy/032C72AEC537EBB6AE96C0FD90387E3E> (дата звернення: 18.02.2026).

194. Chun J., Schroeder de Witt C., Elkins K. Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US : arXiv preprint arXiv:2410.21279. 2024. URL: <https://arxiv.org/abs/2410.21279> (дата звернення: 18.02.2026).

195. Kusche I. Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental Rights and Risk. *Journal of Risk Research*. 2024. DOI: 10.1080/13669877.2024.2350720. URL: <https://www.tandfonline.com/doi/full/10.1080/13669877.2024.2350720> (дата звернення: 18.02.2026).

196. Кожухар О. Г. Правове регулювання систем штучного інтелекту в ЄС: передумови, сучасний стан та перспективи. *Наукові записки НаУКМА. Юридичні науки*. 2024. Т. 13. С. 65–73. DOI: 10.18523/2617-2607.2024.13.65-73. URL: <https://nrplaw.ukma.edu.ua/article/view/313919> (дата звернення: 18.02.2026).

197. Дубняк М. В. Правові підходи в Законі ЄС про штучний інтелект: досвід для України. Інформація і право. 2024. № 3 (50). С. 40–53. DOI: 10.37750/2616-6798.2024.3(50).311600. URL: <http://il.ippi.org.ua/article/view/311600> (дата звернення: 18.02.2026).

198. Nachkevych A. Tools for Adapting Ukraine's Artificial Intelligence Ecosystem to Meet European Union Standards. Право та інноваційне суспільство. 2024. № 1 (22). С. 21–31. DOI: 10.37772/2309-9275-2024-1(22)-2. URL: <https://apir.org.ua/lais/uk/article/view/429> (дата звернення: 18.02.2026).

199. Муравська Ю., Сліпченко Т. Правове регулювання штучного інтелекту в Україні та світі. Актуальні проблеми правознавства. 2024. № 1. С. 188–195. DOI: 10.35774/app2024.01.188. URL: <https://appj.wunu.edu.ua/index.php/apl/article/view/1788> (дата звернення: 18.02.2026).

200. Горобець Н. С., Науменко С. М. Закон ЄС про штучний інтелект і його вплив на судову систему України. Науковий вісник Ужгородського національного університету. Серія «Право». 2025. Вип. 89, ч. 4. DOI: 10.24144/2307-3322.2025.89.4.2. URL: <https://visnyk-pravo.uzhnu.edu.ua/article/view/337005> (дата звернення: 18.02.2026).

201. Олійник О. В. Правове регулювання штучного інтелекту в Україні: виклики та перспективи. Economics, Enterprise, Security. 2025. Vol. 6, no. 4. DOI: 10.70651/3083-6018/2025.6.04. URL: <https://www.eu-scientists.com/index.php/sdel/article/view/268> (дата звернення: 18.02.2026).

202. Bradford A. The Brussels Effect: How the European Union Rules the World. New York : Oxford University Press, 2020. 424 p. DOI: 10.1093/oso/9780190088583.001.0001. URL: <https://academic.oup.com/book/36491> (дата звернення: 18.02.2026).

203. S. and Marper v. the United Kingdom [GC] : applications nos. 30562/04 and 30566/04 : judgment of 4 December 2008. European Court of Human Rights.

ECLI:CE:ECHR:2008:1204JUD003056204. URL: <https://hudoc.echr.coe.int/eng?i=001-90051> (дата звернення: 18.02.2026).

204. Big Brother Watch and Others v. the United Kingdom [GC] : applications nos. 58170/13, 62322/14 and 24960/15 : judgment of 25 May 2021. European Court of Human Rights. ECLI:CE:ECHR:2021:0525JUD005817013. URL: <https://hudoc.echr.coe.int/eng?i=001-210077> (дата звернення: 18.02.2026).

205. Bărbulescu v. Romania [GC] : application no. 61496/08 : judgment of 5 September 2017. European Court of Human Rights. ECLI:CE:ECHR:2017:0905JUD006149608. URL: <https://hudoc.echr.coe.int/eng?i=001-177082> (дата звернення: 18.02.2026).

206. Centrum för rättvisa v. Sweden [GC] : application no. 35252/08 : judgment of 25 May 2021. European Court of Human Rights. ECLI:CE:ECHR:2021:0525JUD003525208. URL: <https://hudoc.echr.coe.int/eng?i=001-210154> (дата звернення: 18.02.2026).

207. Glukhin v. Russia : application no. 11519/20 : judgment of 4 July 2023. European Court of Human Rights. ECLI:CE:ECHR:2023:0704JUD001151920. URL: <https://hudoc.echr.coe.int/eng?i=001-225803> (дата звернення: 18.02.2026).

208. Gaughran v. the United Kingdom : application no. 45245/15 : judgment of 13 February 2020. European Court of Human Rights. ECLI:CE:ECHR:2020:0213JUD004524515. URL: <https://hudoc.echr.coe.int/eng?i=001-200525> (дата звернення: 18.02.2026).

209. Peck v. the United Kingdom : application no. 44647/98 : judgment of 28 January 2003. European Court of Human Rights. ECLI:CE:ECHR:2003:0128JUD004464798. URL: <https://hudoc.echr.coe.int/eng?i=001-60558> (дата звернення: 18.02.2026).

210. Szabó and Vissy v. Hungary : application no. 37138/14 : judgment of 12 January 2016. European Court of Human Rights. ECLI:CE:ECHR:2016:0112JUD003713814. URL: <https://hudoc.echr.coe.int/eng?i=001-160020> (дата звернення: 18.02.2026).

211. Recommendation CM/Rec(2021)8 of the Committee of Ministers to Member States on the Protection of Individuals with Regard to Automatic Processing of Personal Data in the Context of Profiling : adopted 3 November 2021. Council of Europe. URL: <https://search.coe.int/cm?i=0900001680a46147> (дата звернення: 18.02.2026).

212. Recommendation CM/Rec(2022)11 of the Committee of Ministers to Member States on Principles for Media and Communication Governance : adopted 6 April 2022. Council of Europe. URL: <https://search.coe.int/cm?i=0900001680a61712> (дата звернення: 18.02.2026).

213. Recommendation CM/Rec(2022)13 of the Committee of Ministers to Member States on the Impact of Digital Technologies on Freedom of Expression : adopted 28 September 2022. Council of Europe. URL: <https://search.coe.int/cm?i=0900001680a61729> (дата звернення: 18.02.2026).

214. Recommendation CM/Rec(2024)5 of the Committee of Ministers to Member States Regarding the Ethical and Organisational Aspects of the Use of Artificial Intelligence and Related Digital Technologies by Prison and Probation Services : adopted 9 October 2024. Council of Europe. URL: <https://search.coe.int/cm?i=0900001680b1d0e4> (дата звернення: 18.02.2026).

215. Declaration by the Committee of Ministers on the Manipulative Capabilities of Algorithmic Processes : Decl(13/02/2019)1 : adopted 13 February 2019. Council of Europe. URL: <https://search.coe.int/cm?i=090000168092dd4b> (дата звернення: 18.02.2026).

216. Resolution 2341 (2020): Need for Democratic Governance of Artificial Intelligence : adopted 22 October 2020. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/28803/html> (дата звернення: 18.02.2026).

217. Resolution 2342 (2020): Justice by Algorithm – The Role of Artificial Intelligence in Policing and Criminal Justice Systems : adopted 22 October 2020. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/28805/html> (дата звернення: 18.02.2026).

218. Resolution 2343 (2020): Preventing Discrimination Caused by the Use of Artificial Intelligence : adopted 22 October 2020. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/28807/html> (дата звернення: 18.02.2026).

219. Resolution 2344 (2020): The Brain-Computer Interface: New Rights or New Threats to Fundamental Freedoms? : adopted 22 October 2020. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/28810/html> (дата звернення: 18.02.2026).

220. Resolution 2345 (2020): Artificial Intelligence and Labour Markets: Friend or Foe? : adopted 22 October 2020. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/28814/html> (дата звернення: 18.02.2026).

221. Resolution 2485 (2023): Emergence of Lethal Autonomous Weapons Systems (LAWS) and Their Necessary Apprehension Through European Human Rights Law : adopted 12 October 2023. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/31633/html> (дата звернення: 18.02.2026).

222. Opinion 303 (2024): Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law : adopted 18 April 2024. Parliamentary Assembly of the Council of Europe. URL: <https://pace.coe.int/en/files/33517/html> (дата звернення: 18.02.2026).

223. Feasibility Study : CAHAI(2020)23 / Ad Hoc Committee on Artificial Intelligence (CAHAI). Strasbourg : Council of Europe, 2020. URL: <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da> (дата звернення: 18.02.2026).

224. Possible Elements of a Legal Framework on Artificial Intelligence, Based on the Council of Europe's Standards on Human Rights, Democracy and the Rule of Law : CAHAI(2021)08rev / Ad Hoc Committee on Artificial Intelligence (CAHAI). Strasbourg : Council of Europe, 2021. URL: <https://rm.coe.int/cahai-2021-08rev-elements-legal-framework-ai/1680a46c9c> (дата звернення: 18.02.2026).

225. Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law : CAI(2023)01 / Committee on Artificial

Intelligence (CAI). Strasbourg : Council of Europe, 2023. URL: <https://rm.coe.int/cai-2023-01-revised-zero-draft/1680a9a2dd> (дата звернення: 18.02.2026).

226. Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law : CAI(2023)18 / Committee on Artificial Intelligence (CAI). Strasbourg : Council of Europe, 2023. URL: <https://rm.coe.int/cai-2023-18-consolidated-working-draft/1680aba482> (дата звернення: 18.02.2026).

227. Committee on Artificial Intelligence (CAI): Terms of Reference : CM(2021)136-add1final. Strasbourg : Council of Europe, 2021. URL: <https://rm.coe.int/terms-of-reference-of-the-committee-on-artificial-intelligence-for-202/1680a74d2f> (дата звернення: 18.02.2026).

228. European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment : adopted at the 31st plenary meeting of the CEPEJ, Strasbourg, 3–4 December 2018 / European Commission for the Efficiency of Justice (CEPEJ). Council of Europe. URL: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c> (дата звернення: 18.02.2026).

229. Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (ETS No. 108) : 28 January 1981 : as amended by Protocol CETS No. 223, 10 October 2018 (Convention 108+). Council of Europe. URL: <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatyenum=223> (дата звернення: 18.02.2026).

230. Ukraine Signs the Council of Europe Framework Convention on AI. Council of Europe, 2025. URL: <https://www.coe.int/en/web/artificial-intelligence/-/ukraine-signs-the-council-of-europe-framework-convention-on-ai> (дата звернення: 18.02.2026).

231. Reservations and Declarations for Treaty No. 225 – Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225). Council of Europe. URL:

<https://www.coe.int/en/web/conventions/full-list?module=declarations-by-treaty&numSte=225&codeNature=10&codePays=U> (дата звернення: 18.02.2026).

232. Presno Linera M. Á., Meuwese A. Regulating AI from Europe: A Joint Analysis of the AI Act and the Framework Convention on AI. *The Theory and Practice of Legislation*. 2025. Vol. 13, no. 3. P. 292–311. DOI: 10.1080/20508840.2025.2492524. URL: <https://www.tandfonline.com/doi/full/10.1080/20508840.2025.2492524> (дата звернення: 18.02.2026).

233. McGregor L., Murray D., Ng V. International Human Rights Law as a Framework for Algorithmic Accountability. *International and Comparative Law Quarterly*. 2019. Vol. 68, no. 2. P. 309–343. DOI: 10.1017/S0020589319000046. URL: <https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6> (дата звернення: 18.02.2026).

234. Kosta E. Algorithmic State Surveillance: Challenging the Notion of Agency in Human Rights. *Regulation & Governance*. 2022. Vol. 16, no. 1. P. 212–224. DOI: 10.1111/rego.12331. URL: <https://onlinelibrary.wiley.com/doi/10.1111/rego.12331> (дата звернення: 18.02.2026).

235. Szappanyos M. Artificial Intelligence: Is the European Court of Human Rights Prepared? *Acta Humana – Human Rights Publication*. 2023. Vol. 11, no. 1. P. 93–110. URL: https://real.mtak.hu/164151/1/06_szappanyos_93-110_ActaHumana2023_1.pdf (дата звернення: 18.02.2026).

236. Lane L. Clarifying Human Rights Standards Through Artificial Intelligence Initiatives. *International and Comparative Law Quarterly*. 2022. Vol. 71, no. 4. P. 915–944. DOI: 10.1017/S0020589322000380. URL: <https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/clarifying-human-rights-standards-through-artificial-intelligence-initiatives/52D69ACE49CE1E0B5D9E69E51CA14690> (дата звернення: 18.02.2026).

237. Bårdsen A. AI and Human Rights at the European Court of Human Rights. *Verfassungsblog*. 2025. 8 December. DOI: 10.17176/20251208-172230-0. URL:

<https://verfassungsblog.de/ai-and-human-rights-at-the-european-court-of-human-rights/>
(дата звернення: 18.02.2026).

238. Берназюк І. М. Штучний інтелект і права людини: виклики для Європейської конвенції з прав людини. Аналітично-порівняльне правознавство. 2025. Вип. 3, ч. 1. С. 89–98. DOI: 10.24144/2788-6018.2025.03.1.12. URL: <https://journal-app.uzhnu.edu.ua/article/view/332656> (дата звернення: 18.02.2026).

239. ASEAN Guide on AI Governance and Ethics. Version 2. Jakarta : Association of Southeast Asian Nations, 2024. URL: https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf (дата звернення: 18.02.2026).

240. Expanded ASEAN Guide on AI Governance and Ethics – Generative AI. Jakarta : Association of Southeast Asian Nations, 2025. URL: <https://asean.org/wp-content/uploads/2025/01/Expanded-ASEAN-Guide-on-AI-Governance-and-Ethics-Generative-AI.pdf> (дата звернення: 18.02.2026).

241. ASEAN Digital Masterplan 2025. Jakarta : Association of Southeast Asian Nations, 2021. URL: <https://asean.org/wp-content/uploads/2021/08/ASEAN-Digital-Masterplan-2025.pdf> (дата звернення: 18.02.2026).

242. Putra B. A. Governing AI in Southeast Asia: ASEAN's Way Forward. *Frontiers in Artificial Intelligence*. 2024. Vol. 7. Art. 1411838. DOI: 10.3389/frai.2024.1411838. URL: <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1411838/full> (дата звернення: 18.02.2026).

243. Ihnatusha V. Classification of AI systems in international legal regulation: a comparative approach. *Science in the Context of Modern Challenges and Prospects : proceedings of the International Scientific Conference (May 21, 2025)*. Utrecht : Bookmundo, 2025. P. 227–230. DOI: 10.64076/iedc250521.

244. Lu Y. , Tie F. H. A Comparative Analysis of AI Regulation in the EU and ASEAN: The Brussels Effect and Its Implications for Emerging Economies. *Journal of Governance and Regulation*. 2025. Vol. 14, no. 4, special issue. P. 399–413. DOI: 10.22495/jgrv14i4siart16. URL: <https://virtusinterpress.org/A-comparative-analysis-of->

[artificial-intelligence-regulation-in-ASEAN-and-the-European-Union.html](#) (дата звернення: 18.02.2026).

245. Xu J., Lee T., Goggin G. AI Governance in Asia: Policies, Praxis and Approaches. *Communication Research and Practice*. 2024. Vol. 10, no. 3. P. 275–287. DOI: 10.1080/22041451.2024.2391204. URL: <https://www.tandfonline.com/doi/full/10.1080/22041451.2024.2391204> (дата звернення: 18.02.2026).

246. Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity. Addis Ababa : African Union, 2024. URL: <https://au.int/en/documents/20240719/continental-artificial-intelligence-strategy> (дата звернення: 18.02.2026).

247. The Digital Transformation Strategy for Africa (2020–2030). Addis Ababa : African Union, 2020. URL: <https://au.int/en/documents/20200518/digital-transformation-strategy-africa-2020-2030> (дата звернення: 18.02.2026).

248. Unver M. B. Democratic AI Governance: Framing a Vision for Africa in View of the EU Experience. *International Review of Law, Computers & Technology*. 2025. P. 1–23. DOI: 10.1080/13600869.2025.2506924. URL: <https://www.tandfonline.com/doi/full/10.1080/13600869.2025.2506924> (дата звернення: 18.02.2026).

249. Ajuzieogu U. C. From Aspiration to Implementation: The African Union AI Continental Strategy and the Development-Governance Paradox : research paper. African Institute for Artificial Intelligence Policy, 2025. URL: https://www.researchgate.net/publication/396230260_From_Aspiration_to_Implementation_The_African_Union_AI_Continental_Strategy_and_the_Development-Governance_Paradox (дата звернення: 18.02.2026).

250. Santiago Declaration : Ninth Ministerial Conference on the Information Society in Latin America and the Caribbean, Santiago, Chile, 7–8 November 2024. Economic Commission for Latin America and the Caribbean (ECLAC), 2024. URL:

<https://conferenciaelac.cepal.org/9/en/documents/santiago-declaration> (дата звернення: 18.02.2026).

251. Digital Agenda for Latin America and the Caribbean (eLAC2026) : Ninth Ministerial Conference on the Information Society in Latin America and the Caribbean, Santiago, Chile, 7–8 November 2024. Economic Commission for Latin America and the Caribbean (ECLAC), 2024. URL:

<https://conferenciaelac.cepal.org/9/en/documents/digital-agenda-latin-america-and-caribbean-elac2026> (дата звернення: 18.02.2026).

252. The Guiding Manual on the Ethics of Artificial Intelligence Use for GCC Member States. Version 1.0. Riyadh : Secretariat General of the Gulf Cooperation Council, 2023. URL: <https://www.bahrain.bh/wps/wcm/connect/f863157d-2753-4ef7-a4dd-21f8f7c01cdb/>

[The+Guiding+Manual+on+the+Ethics+of+Artificial+Intelligenc+for+GCC+Member+States.pdf?MOD=AJPERES&CVID=psWZnfZ](https://www.bahrain.bh/wps/wcm/connect/f863157d-2753-4ef7-a4dd-21f8f7c01cdb/The+Guiding+Manual+on+the+Ethics+of+Artificial+Intelligenc+for+GCC+Member+States.pdf?MOD=AJPERES&CVID=psWZnfZ) (дата звернення: 18.02.2026).

253. Albous M. R., Al-Jayyousi O. R., Stephens M. AI Governance in the GCC States: A Comparative Analysis of National AI Strategies. Journal of Artificial Intelligence Research. 2025. DOI: 10.1613/jair.1.17619. URL:

<https://jair.org/index.php/jair/article/view/17619> (дата звернення: 18.02.2026).

254. ALECSO Charter on Artificial Intelligence Ethics. Tunis : Arab League Educational, Cultural and Scientific Organization (ALECSO), 2025. URL:

<https://www.alecso.org/publications/wp-content/uploads/2025/08/etic.pdf> (дата звернення: 18.02.2026).

255. Declaration and Plan of Action of the VII Meeting of Ministers and High Authorities of Science and Technology of the OAS "Towards the Safe, Secure, and Trustworthy Development and Deployment of AI in the Americas: The Importance of Governance, Regulatory, and Institutional Frameworks" : CIDI/REMICYT-VII/DEC.1/24 : adopted 13 December 2024, Washington, D.C. Organization of American States. URL:

<https://dig.watch/resource/towards-the-safe-secure-and-trustworthy-development-and->

[deployment-of-ai-in-the-americas-the-importance-of-governance-regulatory-and-institutional-frameworks](#) (дата звернення: 18.02.2026).

256. Strengthening Democracy : OAS General Assembly Resolution AG/RES. 3029 (LIV-O/24) : adopted at the fifty-fourth regular session of the General Assembly, June 2024. Organization of American States. URL: <https://documentsearch.oas.org/legacydocument/en/AG09153E09.docx> (дата звернення: 18.02.2026).

257. Yilma K. Ethics of AI in Africa: Interrogating the Role of Ubuntu and AI Governance Initiatives. *Ethics and Information Technology*. 2025. Vol. 27, no. 2. P. 1–14. DOI: 10.1007/s10676-025-09834-5. URL: <https://link.springer.com/article/10.1007/s10676-025-09834-5> (дата звернення: 18.02.2026).

258. *Blueprint: Artificial Intelligence for Africa*. 1st ed. Kigali : Smart Africa, 2021. URL: https://smartafrica.org/wp-content/uploads/2023/11/70029-eng_ai-for-africa-blueprint-min.pdf (дата звернення: 18.02.2026).

259. Resolution on the Need to Undertake a Study on Human and Peoples' Rights and Artificial Intelligence (AI), Robotics and Other New and Emerging Technologies in Africa : ACHPR/Res.473 (EXT.OS/XXXI) : 25 February 2021. African Commission on Human and Peoples' Rights. URL: <https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-human-and-peoples-rights-and-art> (дата звернення: 18.02.2026).

260. Plantinga P., Shilongo K. et al. Responsible Artificial Intelligence in Africa: Towards Policy Learning. *Data & Policy*. 2024. Vol. 6. DOI: 10.1017/dap.2024.6. URL: <https://www.cambridge.org/core/journals/data-and-policy/article/responsible-artificial-intelligence-in-africa-towards-policy-learning/046FDCF371DF7EB95300BCCB7A887B41> (дата звернення: 18.02.2026).

261. Llorens M. P. Crossing Routes: Artificial Intelligence Governance and Human Rights in Latin America. *CEBRI Journal*. 2023. Year 2, no. 7. P. 81—98. DOI: 10.54827/issn2764-7897.cebri2023.07.03.0X.81-98.en. URL:

<https://cebri.org/revista/en/artigo/108/crossing-routes-artificial-intelligence-governance-and-human-rights-in-latin-america> (дата звернення: 18.02.2026).

262. Africa Declaration on Artificial Intelligence : Kigali, Rwanda, 4 April 2025. URL: <https://c4ir.rw/docs/Africa-Declaration-on-Artificial-Intelligence.pdf> (дата звернення: 18.02.2026).

263. Levy Yeyati E. An Enabling Regulatory Framework for Artificial Intelligence in Latin America and the Caribbean : IDB Technical Note No. IDB-TN-3241. Inter-American Development Bank, 2025. DOI: 10.18235/0013801. URL: <https://publications.iadb.org/en/enabling-regulatory-framework-artificial-intelligence-latin-america-and-caribbean> (дата звернення: 18.02.2026).

264. Supplementary Act A/SA.1/01/10 on Personal Data Protection within ECOWAS : adopted at the Thirty-Seventh Session of the Authority of Heads of State and Government, Abuja, 16 February 2010. Economic Community of West African States. URL: <https://africanlii.org/en/akn/aa-ecowas/act/2010/1-1/eng@2010-12-31> (дата звернення: 18.02.2026).

265. SADC Digital Transformation Strategy and Action Plan. Gaborone : SADC Secretariat, 2023. URL: <https://www.sadc.int/document/sadc-digital-transformation-strategy-sadc-dts> (дата звернення: 18.02.2026).

266. Elmahjub E. Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI. *Philosophy and Technology*. 2023. Vol. 36, no. 4. P. 1—24. DOI: 10.1007/s13347-023-00668-x. URL: <https://link.springer.com/article/10.1007/s13347-023-00668-x> (дата звернення: 18.02.2026).

267. Ali F., Bouzoubaa K., Gelli F. et al. Islamic Ethics and AI: An Evaluation of Existing Approaches to AI Using Trusteeship Ethics. *Philosophy and Technology*. 2025. Vol. 38. Art. 120. DOI: 10.1007/s13347-025-00922-4. URL: <https://link.springer.com/article/10.1007/s13347-025-00922-4> (дата звернення: 18.02.2026).

268. Government AI Readiness Index 2025. Oxford Insights, 2025. URL: <https://oxfordinsights.com/ai-readiness/government-ai-readiness-index-2025/> (дата звернення: 18.02.2026).

269. African Union Convention on Cyber Security and Personal Data Protection : 27 June 2014. African Union. URL: https://au.int/sites/default/files/treaties/29560-treaty-0048_-_african_union_convention_on_cyber_security_and_personal_data_protection_e.pdf (дата звернення: 18.02.2026).

270. Tianjin Declaration of the Council of Heads of State of the Shanghai Cooperation Organisation : 1 September 2025. URL: <https://www.cscoc.org/upload/doc/2025%20tianjin%20declaration%20SCO.pdf> (дата звернення: 18.02.2026).

271. Astana Declaration of the Council of Heads of State of the Shanghai Cooperation Organisation : 4 July 2024. Shanghai Cooperation Organisation. URL: <https://eng.sectscoc.org/20240709/1438929.html> (дата звернення: 18.02.2026).

272. Joint Communiqué of the Twenty-Third Meeting of the Council of Heads of Government of Member States of the Shanghai Cooperation Organisation : 15–16 October 2024. Shanghai Cooperation Organisation. URL: <https://eng.sectscoc.org/20241016/1574921.html> (дата звернення: 18.02.2026).

273. APEC Artificial Intelligence (AI) Initiative (2026–2030) : adopted at the 2025 APEC Leaders' Gyeongju Declaration. Asia-Pacific Economic Cooperation, 2025. URL: [https://www.apec.org/meeting-papers/leaders-declarations/2025/2025-apec-leaders--gyeongju-declaration/apec-artificial-intelligence-\(ai\)-initiative-\(2026-2030\)](https://www.apec.org/meeting-papers/leaders-declarations/2025/2025-apec-leaders--gyeongju-declaration/apec-artificial-intelligence-(ai)-initiative-(2026-2030)) (дата звернення: 18.02.2026).

274. EAC Data Governance Policy Framework : validated at the 3rd East African Community Data Protection Knowledge Exchange & Data Governance Policy Framework Validation Workshop, Kigali, Rwanda, October 2024. East African Community. URL: <https://www.eac.int/press-releases/3195-eac-set-to-advance-data->

[governance-and-protection-with-development-of-a-regional-policy-framework](#) (дата звернення: 18.02.2026).

275. Borislav Tonchev v. Bulgaria : application no. 40519/15 : judgment of 16 April 2024. European Court of Human Rights. URL: <https://hudoc.echr.coe.int/fre?i=001-233106> (дата звернення: 18.02.2026).

276. G7 Hiroshima Leaders' Communiqué : Hiroshima, Japan, 20 May 2023. G7. URL: <https://www.mofa.go.jp/files/100506231.pdf> (дата звернення: 18.02.2026).

277. GPAI and OECD Unite to Advance Coordinated International Efforts for Trustworthy AI : press release, 3 July 2024. OECD. URL: <https://www.oecd.org/en/about/news/speech-statements/2024/07/GPAI-and-OECD-unite-to-advance-coordinated-international-efforts-for-trustworthy-AI.html> (дата звернення: 18.02.2026).

278. GPAI Belgrade Ministerial Declaration : 2024 GPAI Summit, Belgrade, Serbia, 4 December 2024. URL: <https://wp.oecd.ai/app/uploads/2025/01/GPAI-Council-Summary-record-4-Dec-2024.pdf> (дата звернення: 18.02.2026).

279. Halliday T. C., Carruthers B. G. The Recursivity of Law: Global Norm Making and National Lawmaking in the Globalization of Corporate Insolvency Regimes. American Journal of Sociology. 2007. Vol. 112, no. 4. P. 1135–1202. DOI: 10.1086/507855. URL: https://www.researchgate.net/publication/238074301_The_Recursivity_of_Law_Global_Norm_Making_and_National_Lawmaking_in_the_Globalization_of_Corporate_Insolvency_Regimes (дата звернення: 18.02.2026).

280. Оніщенко Н. М. Правове регулювання ШІ-простору: людиноцентристський вимір. Правова держава. 2026. Вип. 37. С. 21–30. DOI: 10.33663/0869-2491-2026-37-21-30. URL: <https://pravova-derzhava.org/rule-of-law-issue-37-2026-21> (дата звернення: 18.02.2026).

281. Block-Lieb S., Halliday T. C. Incrementalisms in Global Lawmaking. Brooklyn Journal of International Law. 2007. Vol. 32, no. 3. P. 851–903. URL:

<https://brooklynworks.brooklaw.edu/cgi/viewcontent.cgi?article=1232&context=bjil>

(дата звернення: 18.02.2026).

282. Hathaway O. A. Between Power and Principle: An Integrated Theory of International Law. The University of Chicago Law Review. 2005. Vol. 72, no. 2. P. 469–536. URL: <https://chicagounbound.uchicago.edu/uclrev/vol72/iss2/2/> (дата звернення: 18.02.2026).

283. Костенко О. В. Metaverse та WEB 4.0: від новітніх технологій до нової філософії та правової реальності : монографія / ДНУ «Інститут інформації, безпеки і права НАПрН України». Київ ; Одеса : Фенікс, 2026. 528 с. ISBN 978-617-8763-14-5.

284. Суржинський М. І. Правотворчість в Україні в епоху штучного інтелекту: виклики та можливості адаптації законодавства до європейських стандартів захисту персональних даних. Правотворчість в Україні в контексті розвитку доктрини і практики міжнародного права : монографія / за ред. Н. М. Оніщенко, Л. Г. Фалалєвої. Київ : Юрінком Інтер, 2025. С. 325–368. ISBN 978-966-667-948-5.

285. The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023 : policy paper / Prime Minister's Office, 10 Downing Street. London, 2023. URL: <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> (дата звернення: 18.02.2026).

286. Ігнатуша В. В. Нормативне закріплення принципу контролю людини над системами ШІ у міжнародно-правових актах. Глобальні виклики та інновації: шляхи розвитку сучасної науки : матеріали X Міжнар. наук. конф. (м. Одеса, 16 січня 2026 р.). Одеса, 2026. С. 235–238. DOI: 10.62731/mcnd-16.01.2026.

287. Ігнатуша В. В. Превентивні інструменти забезпечення контролю людини над системами ШІ у міжнародному праві. The Current State of Development of World Science: Characteristics and Features : proceedings of the X International Scientific and

Theoretical Conference (Lisbon, January 23, 2026). Lisbon, 2026. P. 150–153. DOI: 10.36074/scientia-23.01.2026.

288. Медведєва М. О. Регулювання використання штучного інтелекту в контексті міжнародного права навколишнього середовища. Науковий вісник Ужгородського національного університету. Серія «Право». 2026. Вип. 93, ч. 5. С. 381–387. DOI: 10.24144/2307-3322.2026.93.5.49. URL: <https://visnyk-pravo.uzhnu.edu.ua/article/view/356177> (дата звернення: 14.05.2026).

289. Медведєва М. О. Штучний інтелект і міжнародне право. Європейські перспективи. 2026. № 1. С. 184–194. DOI: 10.71404/EP.2026.1.23. URL: https://ep.unesco-socio.in.ua/wp-content/uploads/archive/EP-2026-1/EP_2026_1_184.pdf (дата звернення: 14.05.2026).

290. Šišková N., Medvedieva M. International Human Rights Law and International Humanitarian Law in the Era of Artificial Intelligence. Актуальні проблеми міжнародних відносин. 2026. Вип. 166, ч. 1. С. 49–59. DOI: 10.17721/apmv.2026.166.1.49-59. URL: <https://apir.iir.edu.ua/index.php/apmv/article/view/4020> (дата звернення: 14.05.2026).

ДОДАТКИ

ДОДАТОК А

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні наукові результати дисертації:

1. Ігнатуша В. В. Розвиток міжнародно-правового регулювання штучного інтелекту в рамках ООН. Аналітично-порівняльне правознавство. 2025. № 3, ч. 3. С. 287–296.

DOI: 10.24144/2788-6018.2025.03.3.45. URL:

<https://journal-app.uzhnu.edu.ua/article/view/335016>

2. Ігнатуша В. В. Автономні системи з ШІ у збройних конфліктах та правове забезпечення дотримання принципу розрізнення. Науковий вісник Ужгородського національного університету. Серія «Право». 2025. Вип. 90, ч. 5. С. 151–161. DOI:

10.24144/2307-3322.2025.90.5.18. URL:

<https://visnyk-pravo.uzhnu.edu.ua/article/view/340844>

3. Ігнатуша В. В. Міжнародно-правові механізми забезпечення контролю людини над розвитком систем штучного інтелекту. Український часопис міжнародного права. 2025. № 2. С. 67–71. DOI: 10.36952/ujil.2025.2.67-71. URL:

<https://jusintergentes.com.ua/images/2025/2/67-71.pdf>

4. Ігнатуша В. В. Правове регулювання принципів прозорості та підзвітності систем ШІ в ЄС: сучасний стан та перспективи розвитку. Юридичний науковий електронний журнал. 2026. № 1. С. 265–272. DOI: 10.32782/2524-0374/2026-1/58.

URL: https://www.lsej.org.ua/1_2026/60.pdf

Наукові праці, які засвідчують апробацію матеріалів дисертації:

5. Ігнатуша В. В. Ризики використання штучного інтелекту в оборонній сфері та способи їх подолання в рамках ЄС та НАТО. Інтеграція України до Європейського

Союзу: виклики в умовах воєнного стану. Пам'яті професора Павла Заворотька та професора Володимира Денисова : зб. матеріалів і тез доп. міжнар. наук.-практ. конф. (м. Київ, 4–6 листопада 2024 р.) / за наук. ред. проф. Л. Г. Фалалєєвої. Київ : Вид-во «Людмила», 2025. С. 299–303. URL: <https://idpnan.org/integration-into-european-union>

6. Ihnatusha V. Classification of AI systems in international legal regulation: a comparative approach. *Science in the Context of Modern Challenges and Prospects : proceedings of the International Scientific Conference (May 21, 2025)*. Utrecht : Bookmundo, 2025. P. 227–230. DOI: 10.64076/iedc250521.

7. Ігнатуша В. В. Нормативне закріплення принципу контролю людини над системами ШІ у міжнародно-правових актах. Глобальні виклики та інновації: шляхи розвитку сучасної науки : матеріали X Міжнар. наук. конф. (м. Одеса, 16 січня 2026 р.). Одеса, 2026. С. 235–238. DOI: 10.62731/mcnd-16.01.2026.

8. Ігнатуша В. В. Превентивні інструменти забезпечення контролю людини над системами ШІ у міжнародному праві. *The Current State of Development of World Science: Characteristics and Features : proceedings of the X International Scientific and Theoretical Conference (Lisbon, January 23, 2026)*. Lisbon, 2026. P. 150–153. DOI: 10.36074/scientia-23.01.2026.