СФЕРА ШТУЧНОГО ІНТЕЛЕКТУ В «М'ЯКОМУ ПРАВІ» ЯПОНІЇ (частина 1 - огляд джерел та тенденції)

ГАЧКЕВИЧ Андрій Олександрович - кандидат юридичних наук, доцент, кафедра міжнародної інформації НУ «Львівська політехніка»

ORCID: 0000-0002-8494-1937

УДК [340.13:004.8](520)

DOI: https://doi.org/10.32782/ep.2024.2.36

У цілях дослідження у правовій системі Японії відібрані акти «м'якого права» для регулювання сфери штучного інтелекту, включаючи вказівки національного рівня від MIAS, METI, FDMA та ін., а також окремі міжнародні документи ОЕСР та «Групи семи» з особливим значенням для цієї держави. Вони були розглянуті нарізно та спільно як система рекомендаційних норм. Результати дослідження представлені у двох частинах, перша з яких присвячена загальному огляду актів та актуальним тенденціям.

Досвід Японії обраний об'єктом для вивчення через те, що вона є однією з держав, вчені та дослідники яких зробили найбільший внесок у виникнення технологій штучного інтелекту, а також світовим лідером з їх впровадження. Японський підхід гнучкого управління до штучного інтелекту викликає великий інтерес з позиції теоретико-методологічного обґрунтування вдосконалення державної політики України в напрямку розвитку штучного інтелекту.

У статті надана характеристика кожного опрацьованого акта «м'якого права». Абсолютна більшість з них вперше введена до наукового обігу в українській правовій науці, як-от: Проєкт Рекомендації зі штучного інтелекту в науково-дослідній діяльності (R&D) 2017 р., Соціальні принципи людиноорієнтованого штучного інтелекту 2019 р., Посібник з використання хмарних сервісів штучного інтелекту 2020 р., Урядові рекомендації зі забезпечення принципів штучного інтелекту 2022 р., та ін.

Особлива увага приділена Проєкту Рекомендацій про штучний інтелект для бізнесу, який у момент написання статті ще перебував на розгляді, а також ініціативам у рамках G7 під головуванням Італії. Серед міжнародних документів розглянуті Рекомендація ОЕСР зі штучного інтелекту та Основи комплексної політики Хіросімського процесу.

У цій статті порушена надзвичайно актуальна проблема управління штучним інтелектом, яка пронизує ряд галузей національного права, зокрема адміністративне. Результати дослідження стануть цінними для осіб, які проявляють зацікавлення в регулюванні сфери штучного інтелекту в Україні та за кордоном, особливо з урахуванням того, що іноземні норми, за винятком ЄС, є маловивченими вітчизняними вченими.

Ключові слова: штучний інтелект, сфера штучного інтелекту, акти «м'якого права» Японії, гнучке управління, регулювання штучного інтелекту в Японії, Суспільство 5.0.

Постановка проблеми

Гнучкий підхід до регулювання (agile governance), для якого характерне широке використання актів «м'якого права», дозволяє замість жорстких нормативних рамок створювати рекомендаційні правила. Такий підхід застосовує Японія для сфери штучного інтелекту, яка відіграє важливе значення в її національній економіці та суспільному розвитку.

Досвід Японії, що почав формуватися в середині 2010-их завдяки діяльності Конференції з розвитку мережевого суспільства штучного інтелекту (далі – Конференції з розвитку) (тут і надалі – авторський пере-

клад назв та положень документів, а також органів влади) та проведенні міжнародних дискусій за участю високорозвинених держав світу «Групи семи» (далі – G7), є цінним джерелом запозичення прогресивних практик та вдосконалення державної політики в напрямку розвитку штучного інтелекту.

Акти «м'якого права» Японії для регулювання сфери штучного інтелекту – і національного, і міжнародного рівня – прийняті за період від 2016 р. до сьогодні, як і проєкти на розгляді, містять ряд цікавих для доктринального опрацювання положень, що стали предметом цього дослідження. Водночас, абсолютна більшість з актів, незважаючи на стрімке зростання інтересу українських учених до правових проблем штучного інтелекту, до цього часу не була введена в обіг вітчизняної науки (за винятком переліку принципів та заходів Міжнародного кодексу поведінки Хіросімського процесу для організацій, які займаються розробками передових систем штучного інтелекту 2023 р. [1, с. 59-60]).

Виходячи з цього, це дослідження з фокусом на актах «м'якого права» Японії присвячене проблемі управління штучним інтелектом. Вона набула багатоаспектного характеру та включає в себе необхідність забезпечення дієвого регулювання для сучасних технологій. З одного боку, таке регулювання повинно сприяти технологічному прогресу, а з іншого – гарантувати дотримання суспільної безпеки та основоположних прав людини.

Стан дослідження проблеми

Потреба в детальнішому вивченні проблем регулювання сфери штучного інтелекту відзначена появою в українській науці останніх років цілої низки праць (Вінникова Н.А., Карчевський М.В., Кривицький Ю.В., Токарєва К. С., Савліва Н. О., Тюря Ю.І. та ін.).

Зростання інтересу до тематики штучного інтелекту в Україні відображає світові тенденції та проявляє себе і через те, що іноземний досвід та дискусії на рівні міжнародних організацій набувають усе більшої важливості (ОЕСР, G7, «Група двадцяти», ООН, ЮНЕСКО, Асоціація держав Південно-Східної Азії, Міжнародна амністія, Всесвітній економічний форум в Давосі тощо).

Хоча окремі питання державної політики Японії в напрямку розвитку штучного інтелекту – особливо стратегічна програма – та формування Суспільства 5.0 були висвітлені в працях Г. Андрощука [2, с. 25], О. Баранова [3, с. 34], Ю. Карпенко [4, с. 40], Ю. Когута [5, с. 91], Ю. Чалюк [6], Л. Требик та М. Зубка [7, с. 551], підхід гнучкого управління до сфери штучного інтелекту та його результати детально не вивчалися. Відповідні акти «м'якого права» Японії лише після опублікування цієї статті будуть введені до наукового обігу в Україні.

При ознайомленні з результатами дослідження може виникнути думка, що «твердого права» для сфери штучного інтелекту в Японії не існує. Така думка не є правильною через існування ряду законів, які містять окремі норми щодо створення та використання технологій штучного інтелекту (про дорожній рух та дорожні транспортні засоби, про захист персональних даних, про прозорість цифрових платформ та ін.) [детальніше про це див: 8]. Водночас, спеціальний закон поки відсутній, а предмет регулювання наявних актів «твердого права» не вказує на сферу штучного інтелекту.

Мета і завдання дослідження

Мета дослідження полягає у комплексному аналізі актів «м'якого права» Японії для регулювання сфери штучного інтелекту.

Результати дослідження можуть бути цінними для теоретико-методологічного обґрунтування вдосконалення державної політики України в напрямку розвитку штучного інтелекту.

Ця стаття відображає ту частину дослідження, яка має відношення до загального огляду актів та актуальних тенденцій. Разом з тим, питання визначення поняття штучного інтелекту, принципів та кола суб'єктів сфери будуть порушені у другій частині.

Для виконання мети у цій статті поставлені такі завдання:

- показати еволюцію рекомендаційних норм для сфери штучного інтелекту, які сформувались у правовій системі Японії за період з 2016 р. до сьогодні;

- узагальнити ознаки актів «м'якого права» Японії, які відображають застосування підходу гнучкого управління та розроблені на основі ідеї балансу між технологічним прогресом та захистом прав людини та суспільної безпеки;
- розглянути основні положення актів «м'якого права» Японії національного та міжнародного рівнів на предмет змістового наповнення сфери штучного інтелекту.

Виклад основного матеріалу

У процесі дослідження автор сформував перелік з дев'яти актів «м'якого права», об'єктом регулювання яких є технології штучного інтелекту.

Варто відзначити, крім вказівок національного рівня від Міністерства внутрішніх справ та зв'язку (Ministry of Internal Affairs and Communications – далі MIAS), Міністерства економіки, торгівлі та промисловості (Ministry of Economy, Trade and Industry – METI), Агентства по боротьбі з пожежами та стихійними лизами (Fire and Disaster Management Agency – FDMA) та ін., опрацьовані декілька міжнародних документів, прийняті під егідою OECP та G7, які займають особливе місце в правовій системі Японії.

Усвідомлюючи, що «м'яке право» саме по собі залишається широко дискутованим явищем, асоційованим переважно з міжнародним правом та неоднозначним в оцінках впливу на правову сферу, ми візьмемо за основу одне з базових розумінь – норми «м'якого права», які містяться в актах рекомендаційного характеру та не мають обов'язкової юридичної сили [9, с. 557]. Також погодимося з тим, що в епоху інформаційного суспільства воно набуває нової цінності, формуючи «дієве нормативне ядро» [10, с. 106-117].

Першою віхою формування системи рекомендаційних правил стала підготовка звіту про роль штучного інтелекту в Мережевому суспільстві мудрості (Wisdom Network Society, WINS) дослідницькою групою, сформованою у складі MIAS влітку 2016 р. Згодом на основі цієї групи Інститут інформаційної та комунікаційної політики започаткував Конференцію з розвитку.

У звіті знаходимо орієнтовний перелік принципів дослідження та розробки штучного інтелекту, який Японія представила на зустрічі міністрів інформації та зв'язку «Групи семи» (далі – G7) у 2016 р. в Такамацу [11, с. 5]. Ці вісім принципів мали першочергове значення при підготовці подальших документів та пояснені в другій частині.

Крім того, у матеріалах відзначені ціннісні ідеї, що повинні лежати в основі розвитку систем штучного інтелекту: повага до людської гідності, індивідуальна автономія, підконтрольність та прозорість системи штучного інтелекту, безпека штучного інтелекту, адекватна реакція на ризики, баланс інтересів, сприяння інноваціям та чесна конкуренція [12, с. 28].

Результати подальшої роботи Конференції з розвитку над поглибленим вивченням принципів були представлені в *Проєкті Рекомендацій зі штучного інтелекту в науково-дослідній діяльності (R&D) для міжнародних дискусій 2017 р.* (далі – Проект ШІ для НДД) [13]. Саме в такій редакції документ набув статусу джерела «м'якого права» та відомий як звід принципів науково-дослідної діяльності для сфери штучного інтелекту (AI R&D Principles).

Варто зауважити, що ціннісний погляд на формулювання правил закладений у збільшенні переваг та одночасному зменшенні ризиків штучного інтелекту.

Важлива функція Проєкту ШІ для НДД полягала в тому, щоб активізувати обговорення створення та використання систем штучного інтелекту як у рамках G7 та OECP, так і за участю суб'єктів сфери на національному та міжнародному рівнях, від яких очікувався обмін досвідом [13, с. 2].

Рекомендації носять загальний характер, тобто підлягають застосуванню до всіх сфер діяльності. Разом з тим, автори висловлюють сподівання щодо появи й спеціальних [13, с. 4]. Як показують результати цього дослідження, через три роки вони були створені – для оцінки надійності у сфері безпеки підприємств та використання хмарних сервісів штучного інтелекту.

Проєкт ШІ для НДД мав на меті сприяти формуванню Мережевого суспільства мудрості, суть якого пояснена в документі

описом ознак: (1) люди живуть у гармонії з системами штучного інтелекту, (2) інформація (зокрема дані) є вільно та безпечно створювана і розповсюджувана, а також систематизована в мережі мудрості, (3) взаємодія між людьми, речами та подіями виходить за просторові межі, (4) творчі та яскраві розробки стають усе більш доступними, (5) виникнення в результаті прогресу мереж зі штучним інтелектом [13, с. 3-4].

Так само у документі визначені концептуальні основи наданих рекомендацій, що дозволяють краще зрозуміти підхід гнучкого управління в цілому, квінтесенцією якого є акти «м'якого права» [13, с. 4-5].

Вони слугують інструментом для розвитку людиноорієнтованого суспільства з наголосом на вже згаданих раніше цінностях людської гідності та індивідуальної автономії. У цьому положенні відображена характерна для Японії гуманістична направленість усіх правил для штучного інтелекту, що повинні відповідати духу прав людини.

Вони підлягають поширенню в якості «м'якого права» серед суб'єктів сфери та втілюють у собі найкращі практики в процесі зростання ролі штучного інтелекту в суспільстві. Те, що такі акти не належать до джерел «жорсткого права», відкриває нові можливості для їхнього вдосконалення та характеризує високий рівень гнучкості, необхідний для динамічної галузі державної політики. Відзначена необхідність популяризації й у міжнародному середовищі, у якому вищезгадана ініціатива Японії з представлення переліку принципів дослідження та розробки штучного інтелекту під час зустрічі міністрів G7 у Такамацу стала чи не першою подібною.

Вони враховують технологічну нейтральність – норми впливають однаково на всі види технологій, без жодних переваг або обмежень для окремих, – як передумови конкуренції та сприятливого становища для розробників штучного інтелекту.

Соціальні принципи людиноорієнтованого штучного інтелекту 2019 р., (далі – Соціальні принципи), на відміну від принципів науково-дослідної діяльності, орієнтовані на забезпечення впровадження технологій штучного інтелекту цілим суспільством. Вони були сформульовані спеціально створеною колегією та схвалені Радою сприяння інтегрованій інноваційній стратегії (Офіс Кабінету міністрів Японії) у взаємозв'язку зі стратегією штучного інтелекту.

Документ містить не тільки сім «соціальних принципів» з конкретизацією їхнього змісту, а й пояснює філософію високотехнологічного Суспільства 5.0 – AI-Ready Society – наступного після інформаційного суспільства сталого розвитку, за якого впроваджені безпечні та ефективні передові технології (штучний інтелект, Інтернет речей, робототехніка та ін.) та створені умови для забезпечення свого благополуччя широким колом людей.

Перехід до Суспільства 5.0 вимагає змін у наступних аспектах:

- *людського потенціалу*, що має відношення до: обізнаності населення з перевагами та недоліками штучного інтелекту, зокрема алгоритмічною упередженістю; можливості людей здійснювати творчу та високопродуктивну діяльність; формування достатньої кількості фахівців у галузі;
- *соціальних систем*, насамперед впровадження штучного інтелекту в таких сферах, як охорона здоров'я, фінанси, страхування, транспорт, енергетика тощо з вдосконаленням відповідних інфраструктур;
- *промисловості* – використання гнучких підходів підприємствами та компаніями, реалізація політики відкритості, а також сприяння проявам творчості та підтримка стартапів;
- *інноваційних систем*, з акцентом на створення R&D екосистеми штучного інтелекту за участю університетів, дослідницьких центрів, сектору промисловості та громадськості, а також забезпечення безперешкодного та захищеного доступу до даних для систем штучного інтелекту;
- *управління:* подальше удосконалення державної політики в напрямку розвитку штучного інтелекту, об'єднання зусиль суб'єктів сфери для виявлення нових проблем, оцінки впливів та визначення подальших регулятивних інструментів [14, с. 5-6].

Крім того, у документі пояснені ключові цінності нового японського Суспільства 5.0, які формують його філософію:

1) людська гідність (штучний інтелект надає інструментарій для покращення життя людей і на фізичному, і на ментальному рівнях);

2) різноманіття та інклюзивність (розвиток штучного сприяє формуванню умов, коли люди є дуже різними за багатьма ознаками, водночас живуть у благополуччі);

3) сталий розвиток (технології штучного інтелекту дозволять знайти нові рішення, усунути соціальну нерівність, зробити успішні кроки до подолання екологічних проблем) [14, с. 4].

В опрацьованому переліку наявні декілька актів міжнародно-правового характеру, включення яких зумовлене тим, що Японія брала активну участь у їхній розробці та схваленні, особливо під час свого головування у 2023 р. у G7 ініціювала обговорення основ комплексної політики. Їхні положення та положення актів «м'якого права» національного рівня від японських адміністративних органів тісно пов'язані між собою та розвивались під взаємним впливом. Для прикладу, Проєкт ШІ для НДД та Соціальні принципи (у редакції проєкту 2018 р.) мали суттєвий вплив на зміст рекомендації ОЕСР у 2019 р.. А правила, розроблені під час Хіросімського процесу, були включені до Проєкту Рекомендацій про штучний інтелект для бізнесу 2024 р.

Рекомендація зі штучного інтелекту Ради ОЕСР від 22 травня 2019 р. (далі – Рекомендація ОЕСР 2019), переглянута 8 листопада 2023 р. для уточнення визначення поняття штучного інтелекту, містить ряд важливих для багатоаспектної проблеми управління штучним інтелектом моментів (у цілях дослідження використано неофіційний переклад [15; 16] – «цей переклад виконаний А. Гачкевичем виключно з інформаційною метою, його точність не може бути гарантована ОЕСР»).

По-перше, використаний *системний підхід* до регулювання сфери штучного інтелекту, на основі якого при вивченні проблеми були взяті до уваги такі явища, як: діяльність багатонаціональних підприємств, поширення цифрових технологій, управління ризиками безпеки, захист приватності та персональних даних, досягнення сталого розвитку та ін.

По-друге, відзначений *революційний потенціал штучного інтелекту*, який може проявляти себе і з позитивної сторони у вигляді покращення добробуту людей і підвищення продуктивності, і з негативної – викликів для забезпечення конкуренції на ринку та прав людини, демократії та цифрової безпеки, а також – проявів дискримінації.

По-третє, підкреслене важливе значення *надійності технологій штучного інтелекту та довіри суспільства до них* у причинно-наслідковому зв'язку, який демонструє доцільність подальшого обговорення сфери штучного інтелекту та ефективної реалізації Принципів відповідального управління надійним штучним інтелектом, які є частиною даної Рекомендації.

По-четверте, наведений перелік *пріоритетів державної політики в напрямку розвитку штучного інтелекту,* що включають:

- інвестиції в дослідження та розробки у сфері штучного інтелекту, у тому числі – у відкриті бази даних;

- вдосконалення цифрової екосистеми для штучного інтелекту, а також впровадження механізмів обміну знаннями та гарантування безпеки даних;

- розвиток інструментів вироблення та реалізації політики, які стимулюють практичне використання розробок після експериментальної перевірки;

- підвищення кадрового потенціалу та підготовка до трансформації ринку праці з врахуванням забезпечення справедливого переходу як системи заходів для пом'якшення негативних наслідків впровадження штучного інтелекту.

По-п'яте, сформовані завдання *міжнародної міжурядової співпраці*: просування схвалених принципів та посилення відповідального управління надійним штучним інтелектом; сприяння обміну знаннями – міжгалузевому і багатосторонньому – та накопичення довгострокового досвіду стосовно штучного інтелекту; прийняття на основі консенсусу технічних стандартів для функціонально сумісних та надійних технологій штучного інтелекту; встановлення порівняльних показників для моніторингу діяльності в сфері штучного інтелекту та оцінювання прогресу в реалізації принципів.

Ще одна розробка Конференції з розвитку – *Рекомендації з використання штучного інтелекту* (далі – Рекомендації з використання) – з'явилася у 2019 р. (проєкт – 2018 р.). Вона певною мірою доповнює Проект ШІ для НДД тими принципами та заходами, які спрямовані на наступні після дослідження та розробки етапи впровадження і застосування (у життєвому циклі технологій штучного інтелекту).

Спорідненість між обома документами зустрічаємо також у наявності концептуальних основ, які для Рекомендацій з використання є наступними: розвиток людиноорієнтованого суспільства та повага до різноманіття (по відношенню до людей, які використовують штучний інтелект); досягнення сталого розвитку разом з вирішенням актуальних проблем; забезпечення балансу при використанні між тими перевагами, які будуть отримані, та ризиками; розподіл ролей між суб'єктами сфери штучного інтелекту; поширення в якості «м'якого права» з усіма відповідними наслідками [17, с. 9].

Рекомендації з використання містять деталізований перелік суб'єктів сфери, який заслуговує окремої уваги у другій частині результатів цього дослідження, а також пов'язаний з ними життєвий цикл технологій штучного інтелекту. Власне положення щодо виділення в процесі створення та використання технологій конкретних завдань для відповідальних суб'єктів, є дуже важливою частиною документу, що підкреслює його цінність та відрізняє від усіх розглянутих вище. Невипадково документ має підназву «Практичний довідник із використання штучного інтелекту».

У цьому дослідженні було опрацьовано два документи спеціального характеру: про застосування штучного інтелекту для хмарних сервісів – *Посібник з хмарних сервісів штучного інтелекту 2020 р.* (далі – Посібник з хмарних сервісів) MIAC [18], для оцінки надійності у сфері безпеки підприємств – *Рекомендації з оцінювання надійності штучного інтелекту у сфері безпеки підприємств 2021 р.* (далі – Рекомендації для безпеки) FDMA [19].

Автори Посібника з хмарних сервісів мали на меті показати, як впроваджувати технології штучного інтелекту постачальникам хмарних сервісів та при цьому підвищувати рівень довіри користувачів. В основу Посібника з хмарних сервісів покладений Проєкт ШІ для НДД.

Пояснені основні етапи розробки та використання технологій штучного інтелекту в хмарних сервісах, а також визначені їхні завдання, а саме: формулювання очікуваних функцій штучного інтелекту та складення плану на *першому етапі планування*; забезпечення засобів для розробки – другому – *підготовки до розробки*; збір і обробка даних, створення моделей, проведення машинного навчання та тестування – *розробки*; налаштування відповідно до потреб користувача – *індивідуального налаштування*; моніторинг та покращення якості – *експлуатації/обслуговування*.

Варто додати, що Посібник з хмарних сервісів також містить рекомендації для постачальників – вони здебільшого полягають у детальному і своєчасному інформуванні користувачів про особливості роботи технологій штучного інтелекту.

Розробка Рекомендацій для безпеки була обумовлена відсутністю методології для оцінювання надійності штучного інтелекту, що значно сповільнювало його впровадження. З одного боку, технології штучного інтелекту мали б сприяти вдосконаленню системи безпеки, у тому числі на підприємствах хімічної та нафтової промисловості, а так само підвищувати їхню ефективність, тоді як з іншого – ці технології могли самі по собі бути не до кінця безпечними, а отже, потребувати гарантій надійності.

Базовим документом для складення рекомендацій стали Вказівки щодо управління якістю машинного навчання [20], які підготував Національний інститут промислових наук і технологій (National Institute of Advanced Industrial Science and Technology, AIST). Підходи до визначення цілей якості та способи їх досягнення взяті за основу для того, щоб показати підприємствам, як покращити рівень безпеки за допомогою високонадійного штучного інтелекту. Водночас, штучний інтелект розглянутий у вузькому розумінні, тобто як технології машинного навчання.

Відповідно до Рекомендацій з безпеки існує три рівні системи управління якістю стосовно штучного інтелекту – внутрішня та зовнішня якість, а також якість у використанні – пов'язані зі штучним інтелектом як технологією в цілому та з окремими складовими (дані, модель, методи впровадження та роботи). Внутрішню якість забезпечує створення складових системи штучного інтелекту з дотриманням відповідних вимог, зовнішню – встановлення необхідних результатів для наявних складових та визначення рівня їхнього досягнення (запобігання ризикам та ефективність) [19, с. 18-26].

Урядові рекомендації із забезпечення принципів штучного інтелекту 2022 р. (далі – Урядові рекомендації) – це інструментарій для компаній та підприємств щодо виконання «соціальних принципів». Вони складаються з комплексу підкріплених прикладами дій при впровадженні штучного інтелекту та описують корпоративну модель гнучкого управління штучним інтелектом, яка складається з п'яти базових процедур [21, с. 9-49]. Вони перебувають у тісній взаємодії та динаміці:

- *аналіз наявних умов та ймовірних ризиків* – прогнозування наслідків впровадження штучного інтелекту, позитивних та негативних, з урахуванням ступеню соціального прийняття використання штучного інтелекту;

- *встановлення цілей управління штучним інтелектом*, зокрема схвалення відповідних корпоративних політик – для сфери штучного інтелекту та пов'язаних із нею сфер (якщо ймовірні ризики є мінімальними, то може бути прийнятим рішення про відмову у створенні системи управління штучним інтелектом);

- *створення системи управління штучним інтелектом*, яке вимагає аналізу розбіжностей між наявним станом та бажаним щодо досягнення встановлених цілей, підвищення рівня фаховості працівників, які задіяні в системі управління, налагодження більш ефективної взаємодії між підрозділами, а також запобігання інцидентам та вчасне реагування на них;

- *впровадження системи управління штучним інтелектом*, що передбачає ведення відповідних записів для загальної характеристики стану впровадження, а також моніторинг роботи окремих систем штучного інтелекту (у подальшому така інформація відіграє важливе значення для вдосконалення системи управління);

- *оцінювання* – перевірка належної роботи системи управління штучним інтелектом, для проведення якої рекомендовано залучати осіб, які не здійснювали розробки штучного інтелекту та не мають посадових обов'язків з використання, а також отримання зворотного зв'язку від зацікавлених сторін (бізнес-партнерів, користувачів, експертів, громадських організацій, професійних спілок).

Під головуванням Японії в 2023 р. у G7 проблемам регулювання сфери штучного інтелекту була приділена особлива увага, а формат обговорення за участю високорозвинених держав світу дозволив зробити значний поступ у справах глобального управління штучним інтелектом.

Два документи – *Міжнародні керівні принципи Хіросімського процесу для організацій, які займаються розробками передових систем штучного інтелекту 2023 р.* (далі – Міжнародні керівні принципи) [22], та *Міжнародний кодекс поведінки Хіросімського процесу для організацій, які займаються розробками передових систем штучного інтелекту 2023 р.* (далі – Міжнародний кодекс поведінки) [23] – стали важливим підсумком обговорення у вигляді формування основ комплексної політики щодо створення та використання передових технологій штучного інтелекту, включаючи генеративний.

Варто зауважити, підписавши відповідну заяву лідерів G7, У. фон дер Ляєн відзначила, що обидва документи уособлюють цінності ЄС щодо розвитку надійного штучного інтелекту та закликала розробників якнайшвидше підписати Міжнародний кодекс поведінки та почати дотримуватись його положень [24].

Основи комплексної політики розроблені на підставі Принципів відповідального управління надійним штучним інтелектом ОЕСР та призначені для діяльності з проєктування, розробки, впровадження та застосування передових систем штучного ін-

телекту всіма суб'єктами сфери штучного інтелекту. Своєю чергою, коло суб'єктів не обмежене приватним сектором, а охоплює також академічні кола, організації громадянського суспільства та державний сектор.

Квінтесенцією обидвох документів є перелік з одинадцяти правил, що підлягають застосуванню на добровільних основах, виходячи з підходу оцінки ризиків, для сприяння безпечному, захищеному та надійному штучному інтелекту. Хоча в назвах зустрічаємо термін «керівні принципи», йдеться радше про керівні вказівки, які слугують забезпеченню принципів у розумінні попередніх документів, у тому числі Соціальних принципів.

1. Для виявлення, оцінювання та пом'якшення ризиків протягом усього життєвого циклу штучного інтелекту вживати відповідні заходи.

2. Визначати та зменшувати слабкі сторони після впровадження технологій, а також мінімізувати інциденти та випадки зловживань за необхідністю.

3. Повідомляти громадськість про можливості, обмеження та сфери належного та неналежного використання систем штучного інтелекту для забезпечення достатньої прозорості та підвищення відповідальності.

4. Сприяти відповідальному обміну інформацією та звітуванню про інциденти за участю розробників передових систем штучного інтелекту, включаючи промисловість, уряди, громадянське суспільство та академічні кола.

5. Розробляти, реалізовувати та оприлюднювати політики для штучного інтелекту та управління ризиками, включаючи політику конфіденційності та заходи з пом'якшення ризиків, зокрема розробникам передових систем штучного інтелекту.

6. Інвестувати та забезпечувати впровадження надійних засобів контролю безпеки, включаючи фізичну безпеку, кібербезпеку та засоби захисту від внутрішніх загроз протягом усього життєвого циклу штучного інтелекту.

7. Розробляти та розгортати надійні механізми автентифікації інформації та встановлення джерела походження, якщо це технічно можливо, наприклад, водяні знаки або інші засоби для визначення користувачам інформації, згенерованої штучним інтелектом.

8. Надавати пріоритет дослідженням щодо зменшення суспільних ризиків безпеки та захищеності, а також надавати пріоритет інвестуванню в ефективні заходи пом'якшення таких ризиків.

9. Надавати пріоритет розробці передових систем штучного інтелекту для вирішення найбільш актуальних викликів у світі, зокрема кліматичної кризи, глобальної охорони здоров'я та освіти.

10. Сприяти розробці та прийняттю за необхідності міжнародних технічних стандартів.

11. Вживати заходів щодо захисту персональних даних та прав інтелектуальної власності при введенні даних.

Актуальні тенденції в «м'якому праві» Японії відображає підготовлений METI та MIAC *Проєкт Рекомендацій про штучний інтелект для бізнесу 2024 р.* (далі – Проєкт для бізнесу), який на момент написання статті пройшов етап громадського обговорення, водночас ще не був схвалений. Проєкт для бізнесу, на думку автора статті, можна поставити в один ряд з такими передовими правовими документами сучасності – «м'якого права» та «твердого права» – як: Закон ЄС про штучний інтелект (AI Act), Проєкт Закону Австралії про штучний інтелект та дані (AIDA), Проєкт Біллю про штучний інтелект та права США (Blueprint for an AI Bill of Rights), а також Типові рамки управління штучним інтелектом Сінгапуру (Model AI Governance Framework), Рекомендація ОЕСР 2019 та обидва документи Хіросімського процесу.

Великою мірою така точка зору обумовлена тим, що Проєкт для бізнесу ввібрав у себе досвід усіх головних попередніх актів «м'якого права» Японії та є спробою систематизації рекомендаційних положень для сфери штучного інтелекту.

Концептуальні основи рекомендацій Проєкту для бізнесу в загальних рисах схожі на визначені для вищерозглянутих актів «м'якого права» Японії національного рівня [13; 14; 17; 21] та полягають у наступному:

1) підтримка добровільних зобов'язань бізнесу щодо виконання рекомендацій зі застосуванням підходу оцінки ризиків;

2) узгодженість з міжнародними дискусіями, які відображають перспективи розвитку сфери штучного інтелекту;

3) забезпечення легкості в розумінні, у тому числі завдяки виокремленню правил для різних категорій суб'єктів (розробників, постачальників та користувачів);

4) багатосторонність процесу створення, завдяки якій рекомендації відображають спільну позицію органів державної влади, академічних кіл, громадянського суспільства та приватного сектора, а отже, матимуть високу ефективність;

5) статус «живого документа», який буде зазнавати змін з розвитком технологій штучного інтелекту [25, с. 4].

Загалом Проєкт для бізнесу складається з п'яти розділів: перші два – терміни та філософсько-ціннісні основи – призначені для всіх суб'єктів бізнесу, а три наступні – для розробників, для постачальників та користувачів поокремо. Зміст документу показує, наскільки змінилось «м'яке право» Японії за майже 10 років – від зводу абстрактних принципів для науково-дослідної діяльності до розгалуженої системи рекомендацій з виділенням різних груп суб'єктів та деталізацією вказівок для кожної з них, а також формуванням ідеології «м'якого права» навколо орієнтирів Суспільства 5.0 та людиноорієнтованого штучного інтелекту.

Перспективним напрямом розвитку регулювання сфери штучного інтелекту є генеративний штучний інтелект, який з появою ChatGPT відкрив людству абсолютно нові можливості. У Японії на цьому етапі обговорюють шляхи та способи обмеження негативного впливу генеративного штучного інтелекту на суспільство за допомогою права, водночас, ні закону, ні спеціальних рекомендацій прийнято ще не було. Разом з тим, на міжнародному рівні питання, ставши предметом дискусій, дало привід для підготовки за участю японських представників ряду документів, наприклад, Звіту Директорату науки, технологій та інновацій ОЕСР «На шляху до загального розуміння генеративного штучного інтелекту в G7» [26], Посібника ЮНЕСКО з генеративного штучного інтелекту в освіті та науці [27] та ін.

Серед інших тенденцій варто відзначити посилення уваги до штучного інтелекту в державному секторі. Декілька тижнів тому (14-15 березня) відбулась зустріч міністрів промисловості, технологій та цифрового розвитку в рамках G7 під головуванням Італії в 2024 р. За її результатами підписана Декларація міністрів, у якій відзначена необхідність пристосування державного сектору до впровадження технологій штучного інтелекту та анонсоване створення відповідного інструментарію найближчим часом, в основу якого покладені принципи безпечного, захищеного та надійного штучного інтелекту. Крім того, підкреслені три пріоритети регулювання в цілях управління штучним інтелектом: підхід оцінки ризиків; відкрите та сприятливе середовище для розробки штучного інтелекту та впровадження; збільшення переваг за зменшення ризиків [28].

Висновки

Внаслідок еволюції рекомендаційних норм, зумовленої швидким прогресом технологій штучного інтелекту, за короткий час – неповне десятиріччя – у Японії сформувалась окрема система «м'якого права».

Результати дослідження показують, що тенденції еволюції визначені рухом від базових принципів абстрактного характеру та цінностей для сфери штучного інтелекту як важливої – можливо найважливішої передумови побудови високорозвиненого в технологічному плані Суспільства 5.0 до деталізованих вказівок для суб'єктів сфери, які за способом вираження нагадують позитивні обов'язки держав та мають на меті захистити суспільну безпеку та основоположні права людини, сприяючи позитивним наслідкам подальшого прогресу технологій штучного інтелекту.

При формуванні системи «м'якого права» на основі рекомендаційних норм, призначених для відносин зі створення технологій штучного інтелекту – розробки та дослідження, – які мають відношенню до найперших етапів їхнього життєвого циклу, з'явились норми щодо впровадження та застосування. Таким чином, сьогодні увесь життєвий цикл отримав правове забезпечення.

Знаковим документом для сфери штучного інтелекту з позиції еволюції є Проєкт для бізнесу, який систематизує рекомендаційні норми та втілює в собі важливі ознаки актів «м'якого права» Японії для сфери штучного інтелекту, зокрема:

- реалізацію стратегічного підходу до побудови Суспільства 5.0;
- наявність концептуальних основ, які характерні й для попередніх документів національного та міжнародного рівнів;
- побудовану на ідеї людиноорієнтованого штучного інтелекту філософсько-ціннісну ідеологію;
- пріоритетність підходу оцінки ризиків для підприємств та компаній, які є суб'єктами сфери штучного інтелекту;

Окремо відзначу добровільність у виконанні, яка, з одного боку, є слабкою стороною «м'якого права» в порівнянні з «твердим правом», з іншого, – призводить до перенесення контролю за діяльністю у сфері штучного інтелекту на підприємства та компанії, а також прийняття ними відповідних політик та впровадження прогресивних практик управління, від яких залежить рівень суспільної довіри та готовність до співпраці партнерів по бізнесу.

Японський підхід гнучкого управління до штучного інтелекту з використанням вищерозглянутих актів «м'якого права» є дуже цінним у контексті вдосконалення державної політики України в напрямку розвитку штучного інтелекту хоча б через те, що Японія відіграє передову роль у міжнародних ініціативах щодо управління штучним інтелектом у рамках G7 та OECP. Крім того, завдяки рекомендаційним нормам у японській правовій системі ми можемо побачити, що у сфері штучного інтелекту відбувається розгалуження суспільних відносин (приклади хмарних сервісів штучного інтелекту та вдосконалення безпеки підприємств за допомогою технологій машинного навчання).

Література

1. Марущак А.І. Вплив міжнародних процесів регулювання штучного інтелекту на інформаційне право України. *Інформація і право*. 2023. №4 (47). С. 55-63. https://doi.org/10.37750/2616-6798.2023.4(47).291582 (дата звернення: 01.04.2024).

2. Андрощук Г.О. Політика і стратегії розвитку штучного інтелекту в країнах світу: quo vadis? (частина 1). *Наука технології інновації*. 2023. №1 (25). С. 15-29.

3. Баранов О.А. Правові аспекти національних стратегій розвитку штучного інтелекту. *Юридична Україна*. 2019. №7. С. 21-38.

4. Карпенко Ю. Зарубіжний досвід визначення концептуальних та інституціональних засад розвитку штучного інтелекту в публічному управлінні. *Теоретичні та прикладні питання державотворення*. 2019. Вип. 24. С. 34-42.

5. Когут Ю. І. Штучний інтелект і безпека: практичний посібник / за ред. проф. А. С. Довгопологого. Київ : Консалтингова компанія «СІДКОН»; ВД Дакор, 2024. 294 с.

6. Чалюк Ю. О. Суспільство 5.0 у японській концепції кейданрен. *Mechanism of an Economic Regulation*. 2023. Вип. 1 (99). С. 65–74. https://doi.org/10.32782/mer.2023.99.11 (дата звернення: 01.04.2024).

7. Требик Л., Зубко М. Використання штучного інтелекту в публічній сфері: зарубіжний досвід. *Current issues of science and integrated technologies* : Proceedings of the First international scientific and practical conference (Milan, January 10-13, 2023). Milan, 2023. Pp. 550-552.

8. Shimpo F. The Principal Japanese AI and Robot Law. Strategy and Research toward Establishing Basic Principles. *Journal of Law and Information System*. 2018. Vol. 3. Pp. 44-65. https://doi.org/10.32235/alis.3.0_44 (дата звернення: 01.04.2024).

9. Мавроматі О. Є. Джерела «м'якого права» в системі права Ради Європи. *Актуальні проблеми держави і права*. 2011. Вип. 62. С. 557-562.

10. Кирилюк О. В. М'яке право як нормативна основа глобального інформаційного суспільства. *Актуальні проблеми міжнародних відносин*. 2015. Вип. 125 (частина І). С. 114-115.

11. AI Networking and the Challenges for the Realization of Wisdom Network Society (WINS) (in Japan). *Ministry of Internal Affairs*

and Communications. URL: https://www.soumu.go.jp/main_content/000425289.pdf (дата звернення: 01.04.2024).

12. AI Networking and the Issues of the Realization of Wisdom Network Society (WINS) (in Japan). *Ministry of Internal Affairs and Communications*. URL: https://www.soumu.go.jp/main_content/000425316.pdf (дата звернення: 01.04.2024).

13. Draft AI R&D Guidelines for International Discussions. *Ministry of Internal Affairs and Communications*.URL: https://www.soumu.go.jp/main_sosiki/joho_tsusin/eng/pressrelease/2017/pdf/000507517.pdf (дата звернення: 01.04.2024).

14. Social Principles of Human-Centric AI. *Cabinet Secretariat*. URL: https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf (датазвернення: 01.04.2024).

15. Recommendation of the Council on Artificial Intelligence (Adopted by the Council at Ministerial Level on 22 May 2019). *OECD Network Environment*. URL: https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf (дата звернення: 01.04.2024).

16. Recommendation of the Council on Artificial Intelligence (Amended on: 08/11/2023). *OECD Legal Instruments*. URL: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449#mainText (дата звернення: 01.04.2024).

17. AI Utilization Guidelines. Practical Reference for AI Utilization. *Ministry of Internal Affairs and Communications*. URL: https://www.soumu.go.jp/main_content/000658284.pdf (дата звернення: 01.04.2024).

18. Guidebook on AI-based Cloud Services (in Japan). *Ministry of Internal Affairs and Communications*. URL: https://www.soumu.go.jp/main_content/000792669.pdf (дата звернення: 01.04.2024).

19. Guidelines on Assessment of AI Reliability in the Field of Plant Safety. *Fire and Disaster Management Agency*. URL: https://www.fdma.go.jp/relocation/neuter/topics/fieldList4_16/pdf/r03/jisyuhoan_shiryo_03_09.pdf (дата звернення: 01.04.2024).

20. Machine Learning Quality Management Guideline. AIST Digital Architecture Research Cente. URL: https://www.digiarc.aist.go.jp/en/publication/aiqm/ (дата звернення: 01.04.2024).

21. Governance Guidelines for Implementation of AI Principles. *Ministry of Economy, Trade and Industry*. URL: https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20220128_2.pdf (дата звернення: 01.04.2024).

22. Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System. *Ministry of Foreign Affairs of Japan*. *URL*: https://www.mofa.go.jp/files/100573471.pdf (дата звернення: 01.04.2024).

23. Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems. *Ministry of Foreign Affairs of Japan*. URL: https://www.mofa.go.jp/files/100573473.pdf (дата звернення: 01.04.2024).

24. Commission welcomes G7 leaders' agreement on Guiding Principles and a Code of Conduct on Artificial Intelligence. European Commission. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5379 (дата звернення: 01.04.2024).

25. (Draft AI Guidelines for Business. *Ministry of Internal Affairs and Communications*. URL: https://www.soumu.go.jp/main_content/000923717.pdf (дата звернення: 01.04.2024).

26. G7 Hiroshima Process on Generative Artificial Intelligence (AI). Towards a G7 Common Understanding on Generative AI. *Organisation for Economic*

Co-operation and Development. URL: https://www.oecd-ilibrary.org/docserver/bf3c0c60-en.pdf?expires=1712619392&id=id&accname=guest&checksum=ABB1AECAFB946C5D34B383FB67D26979 (дата звернення: 01.04.2024).

27. Guidance for generative AI in education and research. UNESCO Digital Library. URL: https://unesdoc.unesco.org/ark:/48223/pf0000386693 (дата звернення: 01.04.2024).

28. G7 Industry, Technology and Digital Ministerial Meeting. Verona and Trento 14-15 March 2024. Ministerial Declaration. *G7 Italia 2024*. URL: https://www.g7italy.it/wp-content/uploads/G7-Industry-Tech-and-Digital-Ministerial-Declaration-Annexes-1.pdf (дата звернення: 01.04.2024).

SUMMARY

For the research, soft law acts on the regulation of AI have been selected in the legal system of Japan. They include domestic guidelines by MIAS, METI, FDMA et al., and having particular importance for Japan international documents of the OECD and the G7. All these acts are examined both separately and together making up the system of recommended norms. This article presents the first part of the results of the study which is devoted to a general overview of AI soft law acts. It also highlights current trends.

Japan was chosen as it is one of the countries whose researchers had made the greatest contribution to the emergence of AI technologies. It belongs to the world leaders in their implementation. At the same time, the Japanese agile governance approach to AI-related issues awakens an interest from the standpoint of theoretical and methodological substantiation of the legal framework for improving the state policy of Ukraine towards the development of artificial intelligence.

This article provides the brief descriptions of selected soft law acts. The absolute majority of them is introduced into scientific circulation of the Ukrainian jurisprudence for the first time. These include the 2017 Draft AI R&D Guidelines for International Discussions, the 2018 Social Principles of Human-Centric AI, the 2020 Guidebook on AI-based Cloud Services which is available only in Japanese, the 2022 Governance Guidelines for Implementation of AI Principles, etc.

The author also examines the 2024 Draft on AI Guidelines for Business pending at the time of writing this article and new Italian initiatives presented at the G7 Ministerial meeting on Industry, Technology, and Digital. The author also discusses the Recommendation of the OECD Council on Artificial Intelligence and the Hiroshima AI Process Comprehensive Policy Framework.

This article raises an extremely urgent problem of governing artificial intelligence which permeates various branches of the national law, in particular administrative law. The results of the study will be of interest to a wide range of scholars who are seeking to understand the complex nature of AI. Relevant issues of foreign AI regulation are poorly studied in Ukraine and little is known about other states than the EU.

Keywords: artificial intelligence, sphere of AI, Japanese soft law acts, agile governance, AI regulation in Japan, Society 5.0.