

УДК 341:004.8

DOI <https://doi.org/10.24144/2788-6018.2024.03.98>

## ХІРОСІМСЬКИЙ ПРОЦЕС ЗІ ШТУЧНОГО ІНТЕЛЕКТУ

**Гачкевич А.О.,***кандидат юридичних наук, доцент,  
доцент кафедри міжнародної інформації  
НУЛП «Львівська політехніка»  
ORCID: 0000-0002-8494-1937*

### **Гачкевич А.О. Хіросімський процес зі штучного інтелекту.**

Обраний предметом вивчення Хіросімський процес зі штучного інтелекту став однією з найважливіших ініціатив глобального управління штучним інтелектом за останні роки. Держави «Групи семи» впродовж 2023 р. проводили обговорення актуальних проблем регулювання сфери штучного інтелекту та прийняли ряд документів, зокрема ті, що формують основи комплексної політики зі штучного інтелекту та рекомендовані до виконання розробниками у різних державах світу.

Автор розглянув Хіросімський процес зі штучного інтелекту як окреме явище з позиції впливу міжнародного співробітництва на розвиток штучного інтелекту. Він опрацював серед іншого Декларацію міністрів з технологій та цифрових питань, Хіросімське комюніке лідерів держав, Міжнародні керівні принципи Хіросімського процесу для організацій, які займаються розробками передових систем штучного інтелекту, а також Звіт «На шляху до загального розуміння генеративного штучного інтелекту в G7», які показують актуальні тенденції управління штучним інтелектом.

Інтерес автора до цього явища великою мірою обумовлений тим, що в процесі обговорення були узгоджені та задокументовані позиції високорозвинених держав, для кожної з яких характерне своє бачення політики та правил для штучного інтелекту. Напрацьовані документи впливатимуть істотним чином на управління штучним інтелектом далеко за межами держав «Групи семи», враховуючи вагому роль останніх у процесах на глобальному рівні.

Внаслідок вивчення Хіросімського процесу зі штучного інтелекту автор виявив, що міжнародне співробітництво за участю «Групи семи» встановлює надбудову для загальноновизнаних принципів надійності, відповідальності, прозорості та справедливості штучного інтелекту. Завдяки аналізу Декларації міністрів сформована свого роду дорожня карта G7 для міжнародних дискусій про передові системи штучного інтелекту. Висновки дослідження можуть бути цінними для кращого розуміння глобального управління

штучним інтелектом, а також для визначення найближчих перспектив регулювання сфери.

**Ключові слова:** штучний інтелект, глобальне управління штучним інтелектом, «Група семи», Хіросімський процес зі штучного інтелекту, Міжнародний кодекс поведінки для розробників, міжнародне співробітництво у галузі штучного інтелекту, дорожня карта для штучного інтелекту, генеративний штучний інтелект, передові системи штучного інтелекту.

### **Hachkevych A. Hiroshima AI Process.**

The focus of this study is on the Hiroshima AI Process. It ranks among the most important AI global governance initiatives in recent years. Over the course of the 2023 year G7 states discussed current issues of AI regulations and adopted a number of documents. Among them are those that lay the foundations of a comprehensive policy framework on AI and are recommended for implementation by the relevant actors in this field.

The author investigated the Hiroshima AI Process as a separate phenomenon from the standpoint of the influence of international cooperation on the development of AI. He examined a few documents reflecting current trends in AI governance. Those include the Takasaki Ministerial Declaration with the results of the meeting of Digital and Tech Ministers, The Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems, as well as the Report "Towards a G7 Common Understanding on Generative AI".

The author's interest in this global governance initiative is largely explained by the fact that the positions of highly developed states with their own vision of AI policy and AI regulations were agreed upon and reflected after the discussions. The documents prepared will have a significant impact on the domain of AI far beyond the borders of G7 states, given their important role in law and policy making processes at the global level.

Subsequently in the study of the Hiroshima AI Process the author revealed that international cooperation with the participation of the G7 had established a superstructure for the generally recognized AI principles of trustworthy,

accountability, transparency and fairness. On the basis of the analysis of the Ministerial Declaration, a kind of AI G7 road map for international discussions on advanced AI systems has been created. The findings are valuable for both a better understanding of AI global governance and determining the near future prospects for regulations.

**Key words:** artificial intelligence, global AI governance, Group of Seven, Hiroshima AI Process, International code of conduct for AI developers, international cooperation in the field of AI, AI road map, generative AI, advanced AI systems.

**Постановка проблеми.** Хіросімський процес зі штучного інтелекту (далі – Хіросімський процес) є цілком новою ініціативою держав «Групи семи» (далі – G7) стосовно проблеми управління штучним інтелектом (далі – ШІ). За короткий час він досягнув успішних результатів внаслідок формування основ комплексної політики щодо створення та використання технологій ШІ у вигляді втілення спільних позицій високорозвинених держав світу.

Як й інші форми міжнародного співробітництва, наприклад, ЄС чи ЮНЕСКО, G7 насамперед завдяки Хіросімському процесу сприяє вдосконаленню інструментів управління та виробленню однотипних підходів до тих або інших питань ШІ. Ця ініціатива також показує перспективи подальшого регулювання сфери, насамперед з приводу пом'якшення ризиків застосування генеративного ШІ.

Проблема управління ШІ, яка сьогодні перебуває в центрі уваги великої кількості вчених сучасності (зокрема українських – Г. Андрощука, Н. Вінникової, Т. Каткової, О. Костенко, І. Онищука та ін.), має багато вимірів: підходи до регулювання у різних державах, напрями розвитку політики в галузі та її тенденції, обов'язки суб'єктів сфери ШІІ (зокрема розробників технологій, постачальників послуг та користувачів систем). Ці та інші виміри перебувають у полі зору й автора цієї статті.

**Мета дослідження.** Головною метою є визначення особливостей Хіросімського процесу як окремого явища міжнародного співробітництва у галузі ШІ. Для її виконання поставлені такі завдання: визначити передумови та ключові події Хіросімського процесу; здійснити аналіз положень основ комплексної політики ШІ, насамперед Міжнародних керівних принципів Хіросімського процесу для організацій, які займаються розробками передових систем штучного інтелекту; охарактеризувати напрями та перспективи управління ШІ, враховуючи позицію Японії в фарватері цього процесу.

**Стан опрацювання проблематики.** Інтерес до Хіросімського процесу з боку українських

вчених помітний у ряді статей за останні місяці, кількість яких швидше за все надалі зростатиме. Згадки про нього зустрічаємо у таких авторів, як І. Басиста, Ж. Удовенко, М.-М. Кулинич (започаткування в контексті огляду тенденцій щодо регулювання ШІ [1, с. 27]), А. Марущак (наведення основних положень одного з найважливіших актів Хіросімського процесу [2, с. 59-60]), О. Подільчак (відзначення як доповнення до юридично обов'язкових правил регулювання у сфері інформаційно-комунікаційних технологій [3, с. 572]). Разом з тим, до цього часу Хіросімський процес не був систематично й ґрунтовно розглянутий як окреме явище, хоча й належить до найважливіших ініціатив глобального управління ШІ сьогодення. Подібне становище склалось і в іноземній літературі, що частково можна пояснити фактором часу. Окремі ідеї щодо ролі Хіросімського процесу в розвитку наукового базису управління ШІ можемо відзначити в працях таких авторів: І. Бреммер та М. Сулейман [4], К. Чан [5], група дослідників з Інституту ініціатив майбутнього Токійського університету [6] тощо.

**Виклад основного матеріалу.** Хіросімським процесом в даній статті позначена ініціатива з узгодження позицій держав G7 щодо управління передовими системами ШІ для збільшення переваг їхнього використання та пом'якшення ризиків. Вона пов'язана з проведенням переговорів та прийняттям домовленостей високорозвиненими державами світу протягом 2023 р.

Не випадковим стало те, що процес відбувався у рік головування Японії в G7, яка, по-перше, належить до світових лідерів з розробки та впровадження технологій ШІ, по-друге, ще в 2016 р., під час попереднього головування, вона заявила про свій намір докласти зусиль та сприяти якнайшвидшому формуванню глобальних правил для ШІ [7].

Проведення саміту G7 у травні 2023 р. та прийняття комюніке за його підсумками, яке вплинуло на подальший хід Хіросімського процесу, саме в цьому місті було символічним [7, с. 2]. Одним з найважливіших питань для Японії як держави, що головувала, перед усіма іншими називалось ядерне роззброєння. Даний саміт мав важливе значення і для Української держави через те, що в ньому взяв участь Президент України В. Зеленський, а учасники заявили про підтримку у боротьбі з агресором.

В період з квітня до листопада проведені ряд зустрічей на найвищих рівнях та представлені важливі документи, в тому числі з основ комплексної політики ШІ, які отримали схвалення глав держав G7.

В основі Хіросімського процесу лежать три засадничі ідеї.

1. Баланс між інноваціями та технологічним прогресом, з одного боку, а з іншого – за-

хистом прав людини та безпеки суспільства. В ширшому контексті сутність балансу полягає в узгодженні усталеного та нового.

2. Орієнтованість на багатостороннє партнерство для вдосконалення управління ШІ як на рівні міжнародних ініціатив, так і через залучення різних категорій зацікавлених сторін (від приватного сектору до громадськості).

Ідея партнерства пов'язана з міжнародним співробітництвом не тільки між державами G7, а й за участю їхніх союзників у вдосконаленні системи глобального управління, зокрема таких міжнародних організацій, як ООН, ЮНЕСКО та ОЕСР, а також ініціативи Глобального партнерства зі штучного інтелекту, започаткованої завдяки попередній діяльності G7.

3. Першочерговість обговорення передових систем ШІ, які в умовах відсутності термінологічної чіткості часто асоціюються з генеративним ШІ. Разом з тим, більш точним є твердження про використання в них методик глибокого навчання та приналежність до найбільш прогресивних розробок. Слід відзначити, що понятійні питання вимагають подальшого вивчення, як і дослідження правильних відповідників для англійських термінів, використовуваних в тому числі у документах G7.

Важливими передумовами Хіросімського процесу стали: технічні, пов'язані з прогресом у технологіях та відображені факторами – 1-3, політико-правові, пов'язані з правовою сферою та державною політикою – 4-6, а також безпекові, які стосуються викликів та ризиків ШІ для людини та суспільства, – 7-9:

1) розширення впровадження систем ШІ у різні сфери діяльності, зокрема промисловість, освіту, охорону здоров'я, торгівлю, управління та ін.;

2) розвиток робототехніки (зокрема у військовій сфері та автомобільній промисловості) та збільшення виробництва промислових роботів;

3) набуття поширення Інтернетом речей, квінтесенцією якого на макрорівні стають «розумні міста»;

4) посилення визнання загальноприйнятих принципів ШІ, до яких належать: надійність, відповідальність, прозорість та справедливість;

5) прийняття та оновлення національних стратегій ШІ для визначення пріоритетів розвитку галузі у державах світу;

6) формування нового законодавства про ШІ та прийняття актів «м'якого права» (наприклад, Закону ЄС про штучний інтелект, Типових рамок управління штучним інтелектом Сінгапуру);

7) зростання актуальності підходу оцінки ризиків, який потребує спеціальної методології;

8) зміщення фокусу уваги з функціональних можливостей систем ШІ (що виконує) на операційний процес (як виконує) через небезпечну невідомість чорного ящика;

9) фіксування великої кількості випадків порушень прав інтелектуальної власності (на тлі прогалин у відповідному законодавстві), а також виникнення загроз для конфіденційності.

Водночас, найбільш суттєвою передумовою, яка впливає на всі перелічені вище, є створення та запуск восени 2022 р. ChatGPT та подальше поширення технологій генеративного ШІ. Також варто особливо відзначити збільшення внеску ШІ у зростання економіки.

*Основні події.* Вперше питання ШІ за головування Японії у 2023 р. було порушеним під час зустрічі міністрів з цифрових питань та технологій у м. Такасаки у квітні 2023 р. Японію представляли три міністри: цифрової трансформації, внутрішніх справ та комунікацій, а також економіки, торгівлі та промисловості. Декларація міністрів від 30 квітня 2023 р. за результатами зустрічі визначила пріоритетні виклики цифрового суспільства:

1) спрощення транскордонних потоків даних на основі правила data free flow with trust («вільний потік даних з довірою»);

2) розвиток цифрової інфраструктури для сталого розвитку з врахуванням безпекових викликів;

3) вдосконалення управління Інтернетом, орієнтоване на запобігання фрагментації;

4) впровадження інноваційних технологій (цифрова ідентифікація, належне програмне забезпечення для цифрових товарів, зменшення негативного впливу фізичних об'єктів на навколишнє середовище, метавсесвіт та ін.);

5) цифрова конкуренція з наголосом на важливості створення сприятливих для інновацій умов;

6) утвердження відповідального ШІ та підтримка глобального управління [8].

В положеннях про ШІ (п. 39–48) згадані інші міжнародні ініціативи, насамперед під егідою ОЕСР (Рекомендація зі штучного інтелекту Ради ОЕСР від 22 травня 2019 р.); названі підходи, методи та інструменти для управління ШІ, зокрема підхід оцінки ризиків та метод рекомендацій; підкреслені цінності, які так само лежать в основі Принципів відповідального управління надійним штучним інтелектом ОЕСР (далі – принципів ОЕСР); відзначена особлива роль генеративного ШІ з усіма його можливостями та загрозами.

Зміст цих положень систематизований на рис. 1, який автор статті вважає свого роду дорожньою картою для міжнародних дискусій про передові системи ШІ на даному етапі.

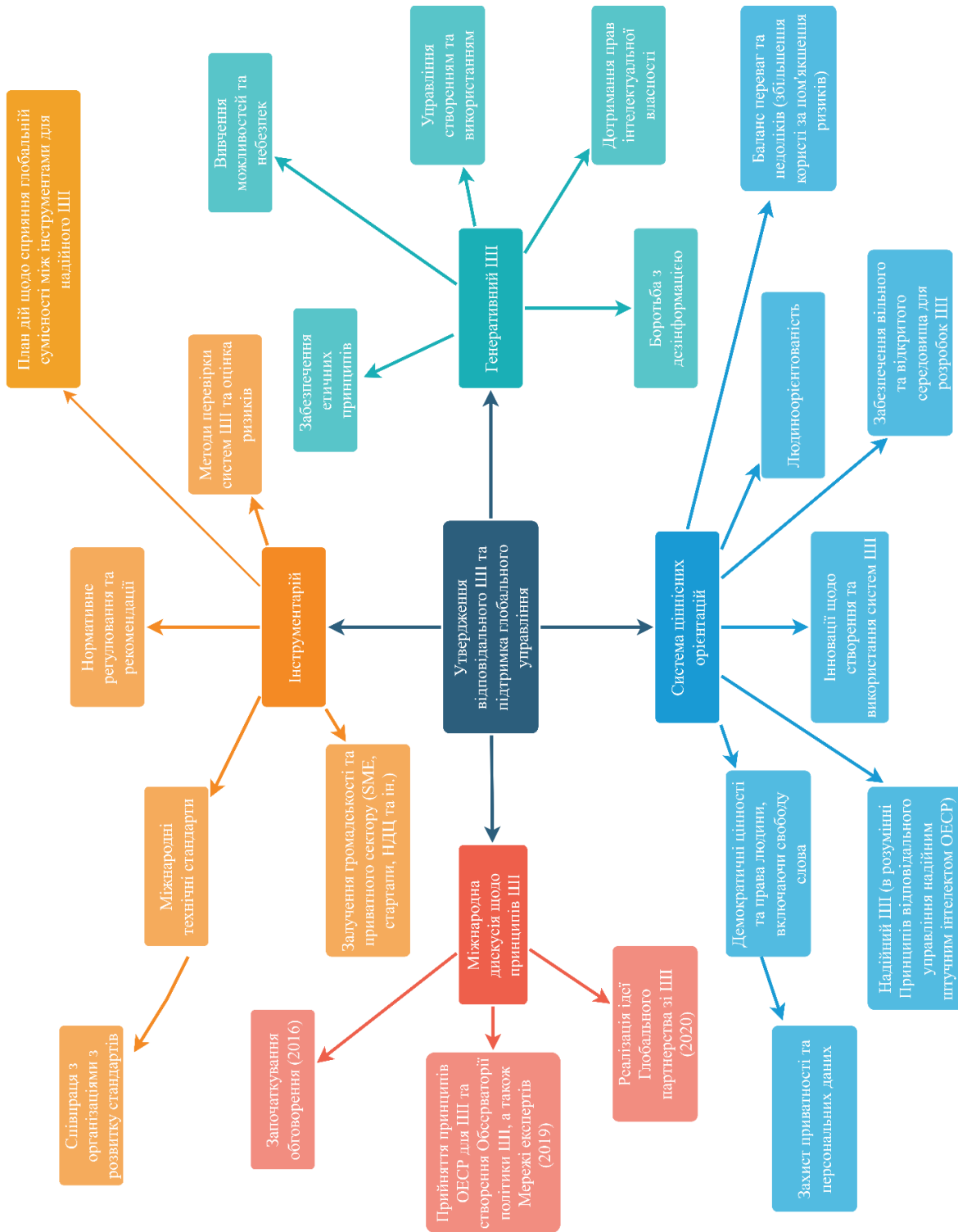


Рис. 1. Дорожня карта G7 для міжнародних дискусій про передові системи ШІ (опрацювання автора на основі 8)

На карті показаний План дій G7 для сприяння глобальній сумісності між інструментами для надійного штучного інтелекту, який є додатком до Декларації. В ньому передбачено наступні напрями для співробітництва щодо забезпечення сумісності рамок управління ШІ: розвиток міжнародних технічних стандартів, проведення воркшопів та круглих столів за участю експертів та дослідників для усестороннього обговорення способів підвищення надійності ШІ, поглиблення участі в міжнародних дискусіях, а також підтримка зусиль для забезпечення демократичних цінностей та поваги до прав людини, взявши за орієнтир принципи ОЕСР у державах з економікою, що розвивається [9].

В Хіросімському комюніке у травні 2023 р. глави держав виділили як окреме завдання сприяння міжнародним дискусіям щодо інклюзивного управління ШІ та сумісності для реалізації їхнього спільного бачення та досягнення мети надійного ШІ [10, с. 2]. Більш детально перспективи співпраці наведені у п. 38, який серед іншого ілюструє одну з правових дилем сучасності: технологічний розвиток випереджає темпи міжнародного управління новими цифровими технологіями, в тому числі ШІ [10, с. 27]. При цьому в документі продемонстрована відданість цінностям справедливості, відповідальності, прозорості, безпеки, поваги до приватного життя та прав людини, захисту персональних даних та ін.

В загальному положення стосовно ШІ п. 38 Комюніке відтворюють домовленості міністрів з цифрових питань та технологій від 30 квітня 2023 р. Водночас, серед них знаходимо доручення зі створення робочої групи для вивчення показаної на рис. 1 проблематики генеративного ШІ у взаємодії з ОЕСР та Глобальним партнерством зі штучного інтелекту (відповідний звіт охарактеризований далі у статті). У тексті також відзначено започаткування Хіросімського процесу зі ШІ [10, с. 28].

Наступна подія відбулась на початку вересня у форматі онлайн для того, щоб узагальнити результати роботи попередніх місяців. Зустріч міністрів з цифрових питань та технологій, а також їх партнерів завершилась прийняттям відповідної заяви та наступними домовленостями:

- визнання розуміння головних ризиків, викликів та можливостей генеративного ШІ відповідно до підготовленого ОЕСР та робочою групою звіту (додаток до Заяви);
- взяття на себе зобов'язання сформуванню керівні принципи та міжнародний кодекс поведінки для організацій, що розробляють передові системи ШІ (приклади принципів наведені);
- підтвердження наміру розвивати проектно співпрацю, яка зокрема має бути зосередженою на технічних можливостях для розріз-

нення недостовірної інформації та дезінформації, включаючи іноземну маніпулятивну інформацію, за допомогою ШІ [11].

Важливою частиною Заяви є пояснення основ комплексної політики, які стали одним з найбільш досягнень процесів глобального управління ШІ:

- а) підтримка інноваційності відповідального ШІ;
- б) сприяння розвитку режимів для регулювання та управління, які узгоджені з національними підходами;
- в) залучення та охоплення зацікавлених сторін;
- г) статус «живого документа», який підлягає оновленням та доповненням;
- д) вдосконалення та реалізація суб'єктами повсюди у світі, включаючи держави з економікою, що розвивається [11].

Наприкінці жовтня 2023 р. Хіросімський процес досягає свого апогею внаслідок схвалення двох знакових документів – Міжнародних керівних принципів для організацій, які розробляють передові системи штучного інтелекту (далі – Міжнародних керівних принципів) та Міжнародного кодексу поведінки для організацій, які розробляють передові системи штучного інтелекту (далі – Міжнародного кодексу).

В Міжнародних керівних принципів перелічені одинадцять заходів для усіх організацій – не тільки з приватного сектору, – рекомендованих до виконання протягом життєвого циклу технологій ШІ (рис. 2).

Акцент щодо об'єкта застосування зроблений на передових системах, серед яких названі два види: генеративні та найбільш прогресивні базові. Перелік є невичерпним та може зазнавати змін. Він був сформований у відповідності до принципів ОЕСР. Відзначена можливість держав вибирати підхід та способи реалізації даних принципів. Вказані обмеження для інновацій, які пов'язані з заборонаю розробляти або розгортати окремі передові системи ШІ (зокрема через такі причини: суперечність демократичним цінностям, заподіяння істотної шкоди для окремих осіб або громад, сприяння тероризму, уможливлення злочинних зловживань або створення серйозних ризиків для безпеки, захисту та прав людини) [12, с. 1-2]. Той самий перелік використаний в Міжнародному кодексі – лише з більшою деталізацією положень [13]. Варто зауважити, що дванадцятье правило – заохочувати та сприяти надійному та відповідальному використанню передових систем ШІ – доповнює перелік Міжнародних керівних принципів для всіх суб'єктів штучного інтелекту. Документ з відповідною назвою доданий до Заяви міністрів від 1 грудня [14, с. 5-6].

Міжнародні керівні принципи	Пункт 1. Вживати відповідні заходи при розробці для виявлення, оцінювання та пом'якшення ризиків протягом усього життєвого циклу ШІ.
	Пункт 2. Визначати та зменшувати слабкі сторони після впровадження технологій, а також знижувати кількість інцидентів та випадків зловживань за необхідністю.
	Пункт 3. Повідомляти громадськість про можливості, обмеження та сфери належного та неналежного використання систем ШІ для забезпечення достатньої прозорості та підвищення відповідальності.
	Пункт 4. Сприяти відповідальному обміну інформацією та звітуванню про інциденти за участю розробників передових систем ШІ (промисловість, уряди, громадянське суспільство та академічні кола).
	Пункт 5. Розробляти, реалізовувати та оприлюднювати політики для штучного інтелекту та для управління ризиками, включаючи політику конфіденційності та заходи з пом'якшення ризиків.
	Пункт 6. Інвестувати та забезпечувати впровадження надійних засобів контролю безпеки, включаючи фізичну безпеку, кібербезпеку та засоби захисту від внутрішніх загроз протягом життєвого циклу ШІ.
	Пункт 7. Розробляти та розгортати надійні механізми автентифікації інформації та встановлення джерела походження по можливості для визначення користувачам інформації, згенерованої ШІ.
	Пункт 8. Надавати пріоритет дослідженням щодо зменшення суспільних ризиків безпеки та захисту, а також надавати пріоритет інвестуванню в ефективні заходи пом'якшення таких ризиків.
	Пункт 9. Надавати пріоритет розробці передових систем ШІ для вирішення найбільш актуальних викликів.
	Пункт 10. Сприяти розробці та прийняттю за необхідності міжнародних технічних стандартів.
	Пункт 11. Вживати заходів із захисту персональних даних та прав інтелектуальної власності при введенні даних.

Рис. 2. Перелік рекомендованих заходів за Міжнародними керівними принципами та Міжнародним кодексом (опрацювання автора на підставі 12, 13)

Прийняття Міжнародних керівних принципів та Міжнародного кодексу відзначене разом зі закликом виконувати їхні положення в Заяві лідерів G7 від 30 жовтня 2023 р. Відзначена необхідність подальшої розробки комплексної політики для систем ШІ під керівництвом міністрів, об'єднуючи зусилля з Глобальним партнерством зі штучного інтелекту та ОЕСР, а також проведення консультацій з усіма зацікавленими сто-

ронами навіть за межами G7 – урядами, академічними колами, громадянським суспільством та приватним сектором [15].

Заява міністрів від 1 грудня 2023 р. прийнята з метою схвалення основ комплексної політики. Їх названо «першою успішною міжнародною базою, що складається з керівних принципів та кодексу поведінки стосовно впливу передових систем ШІ на наші суспільства та економіки»

(за Заявою основи мають чотири складові: 12, 13, 17 та проектна співпраця) [14, с. 1]. У документі також наявний План сприяння Хіросімському процесу, в якому названі такі напрями діяльності: розширення урядової підтримки та заохочення виконання Міжнародних керівних принципів для всіх суб'єктів штучного інтелекту та Міжнародного кодексу, включаючи розробку інструментів моніторингу та механізмів звітності; продовження співробітництва щодо виконання проектів за участю Глобального партнерства зі штучного інтелекту, ОЕСР, ЮНЕСКО, зокрема для вивчення заходів та практик протидії дезінформуванню, забезпечення прозорості та вирішення інших проблем, пов'язаних з генеративним ШІ; створення спеціального веб-сайту для оприлюднення інформації про оновлення політики у державах G7 та список організацій, які взяли на себе зобов'язання виконувати Міжнародний кодекс [14, с. 2].

Підсумковим документом головування Японії у G7 (хронологія подій, пов'язаних з обговоренням управління ШІ, показана у таб. 1) стала Заява лідерів від 6 грудня 2023 р. У одному з розділів – «Цифрова сфера» – глави держав відзначили важливість двох питань для міжнародних дискусій: інклюзивного управління ШІ та сумісності рамок управління ШІ. Крім того, названа спільна мета для держав G7 в умовах наявності різних підходів та інструментів політики у них поокремо, яка полягає у розвитку надійного, безпечного та захищеного ШІ, що відповідає демократичним цінностям. Заява демонструє підтримку основ комплексної політи-

ки та представленого плану, а також черговий раз закликає суб'єктів сфери імплементувати Міжнародні керівні принципи та Міжнародний кодекс [16, с. 13-14].

*Генеративний ШІ.* Особлива роль Хіросімського процесу полягає не тільки у прийнятих рекомендаційних правилах для суб'єктів ШІ, а й представленому звіті за результатами вивчення ряду аспектів, пов'язаних з генеративним ШІ: тенденції розвитку на основі даних Обсерваторії ОЕСР зі ШІ; оцінювання можливостей, ризиків, пріоритетів та заходів впливу; політики та засобів регулювання у державах G7.

Звіт підготовлений Директоратом науки, технологій та інновацій ОЕСР та названий «На шляху до загального розуміння генеративного штучного інтелекту в G7». Він систематизує важливу інформацію для підготовки до обговорень в межах Хіросімського процесу проблематики генеративного ШІ як пріоритетної за Заявою лідерів від 21 травня 2023 р.

В цілому документ має важливе доктринальне значення для управління генеративним ШІ, який часто асоціюється з використанням передових систем (advanced systems), хоча й не зводиться до них. В основі звіту лежить опитування представників держав G7. Завдяки опитуванню можна скласти уявлення про їхнє бачення різних аспектів, які мають визначальний вплив на державну та міжнародну політику в галузі генеративного ШІ. В таб. 2 наведені варіанти вибору, які узагальнюють фактори вдосконалення управління ШІ.

Звіт також показує деякі особливості систем генеративного ШІ, які призводять до появи

Дата події	Назва події та підсумковий документ
29-30 квітня	Зустріч міністрів з цифрових питань та технологій, Декларація міністрів з цифрових питань та технологій
19-21 травня	49-й саміт G7, Хіросімське комюніке лідерів держав
липень – серпень	Підготовка звіту «На шляху до загального розуміння генеративного штучного інтелекту в G7» ОЕСР
7 вересня	Зустріч міністрів з цифрових питань та технологій, а також їх партнерів (ОЕСР, Глобальне партнерство зі ШІ), Заява міністрів з цифрових питань та технологій Хіросімського процесу зі штучного інтелекту
30 жовтня	Зустріч глав держав, Заява лідерів G7 щодо Хіросімського процесу, Міжнародні керівні принципи та Міжнародний кодекс
1 грудня	Зустріч міністрів з цифрових питань та технологій, а також їх партнерів (ОЕСР, Глобальне партнерство зі штучного інтелекту), Заява міністрів з цифрових питань та технологій Хіросімського процесу зі штучного інтелекту
6 грудня	Зустріч глав держав, Заява лідерів G7

Таб. 1. Хронологія Хіросімського процесу (опрацювання автора)

Можливості	Ризики	Пріоритети	Заходи впливу
<ul style="list-style-type: none"> <li>боротьба зі злочинністю та кіберзахист,</li> <li>викорінення бідності та голоду,</li> <li>підвищення продуктивності,</li> <li>розширення персоналізованої освіти для всіх,</li> <li>підтримка людей похилого віку,</li> <li>розвиток сфери розваг та креативного сектору,</li> <li>сприяння вирішенню кліматичної кризи,</li> <li>сприяння інноваціям і підприємництву,</li> <li>удосконалення охорони здоров'я.</li> </ul>	<ul style="list-style-type: none"> <li>зникнення людського фактору,</li> <li>довготермінові екзистенційні загрози,</li> <li>загрози кіберзахисту,</li> <li>загрози конфіденційності,</li> <li>загрози незаконної діяльності (шахрайство в Інтернеті),</li> <li>негативний вплив на навколишнє середовище,</li> <li>переворот на ринку праці,</li> <li>поглиблення соціального розшарування,</li> <li>порушення прав інтелектуальної власності,</li> <li>поширення дезінформації та маніпулювання,</li> <li>розширення нерівності.</li> </ul>	<ul style="list-style-type: none"> <li>безпека людей,</li> <li>вплив на ринки праці, податкові наслідки,</li> <li>демократичні цінності,</li> <li>захищеність та надійність системи,</li> <li>змінювані потреби навчання,</li> <li>інклюзивність та благополуччя,</li> <li>конфіденційність та управління даними,</li> <li>людський фактор проти маніпулювання,</li> <li>основоположні права людини,</li> <li>придатність до пояснення,</li> <li>справедливість та неупередженість,</li> <li>сталий розвиток та використання енергії,</li> <li>узгодженість систем штучного інтелекту та людських цінностей.</li> </ul>	<ul style="list-style-type: none"> <li>добровільні кодекси поведінки,</li> <li>залучення до обговорення,</li> <li>залучення до регулятивних експериментів,</li> <li>механізми впровадження, як-от: стандарти, сертифікації, відшкодування шкоди та ін.,</li> <li>моніторинг інцидентів,</li> <li>оцінювання розвитку подій за допомогою прогнозування,</li> <li>створення ефективних інструментів для безпеки, контролю якості, продуктивності та зміцненню довіри,</li> <li>стимулювання оцінювання ризиків,</li> <li>стимулювання подолання ризиків,</li> <li>стимулювання розробки інструментів для ШІ.</li> </ul>

Таб. 2. Можливості [17, с. 13], ризики [17, с. 14], пріоритети [17, с. 17] та заходи впливу [17, с. 26] для генеративного ШІ (варіанти вибору респондентами)

труднощів з приводу ефективного регулювання сфери:

– непередбачуваність, адаптивність, автономність та багатоцільове призначення, які можуть спричинити такі наслідки, як підривні інновації, незрозумілість логіки роботи системи, посилення упереджень та ін.;

– брак прозорості, відстежуваний під час створення та використання, для виправлення якого потрібні дієві та прогресивні правила не тільки «м'якого права», а й «твердого»;

– можливості розповсюдження недостовірної інформації та дезінформування, що дуже швидко розвиваються та можуть набувати негативно впливу на поведінку людей;

– робота на основі великих обсягів інформації, які своєю чергою можуть бути отриманими з порушенням прав, наприклад, стосовно конфіденційності та інтелектуальної власності [17, с. 22-23].

Загалом матеріали містять наступні відомості: відзначені пріоритети політики в галузі генеративного ШІ, оцінені позитивні переваги та небезпеки; показані національні та регіональні (ЄС) ініціативи для регулювання, які стають

все більш поширеними; висвітлені перспективи розвитку явища, які стали особливо помітними починаючи від осені 2022 р.; наведені засадничі ідеї для формування міжнародного підходу до генеративного ШІ та ін.

**Висновки.** Узагальнюючи результати дослідження, можна зробити висновок, що Хіросімський процес як міжнародна ініціатива покликаний сприяти формуванню глобальної системи управління ШІ. Її суть полягає у виробленні та втіленні однотипних підходів та створенні загальних стандартів по відношенню до регулювання сфери ШІ. Водночас, така система, що є плодом зусиль здебільшого G7, ОЕСР, принципи якої виступають базисом, а також Глобального партнерства зі штучного інтелекту, не є універсальною, хоча й об'єднує високорозвинені держави. У зв'язку з цим, існує необхідність пошуку спільного знаменника з іншими подібними системами або навіть окремим державами, які створюють та використовують технології ШІ. Ця необхідність призвела до пріоритетності сумісності як важливої ідеї Хіросімського процесу. Ще одна ідея, яка радше має відношення до управління на національному рівні, пов'язана з впровад-



женням підходу оцінки ризиків та здійсненням ефективного управління ризиками усіма суб'єктами сфери. Напрацьовані документи засвідчили відданість держав G7 принципам безпечного, надійного та відповідального ШІ, який повинен відповідати демократичним цінностям та сприяти дотриманню прав людини. Значущість Хіросімацького процесу підкреслюють запропоновані основи комплексної політики, частиною яких є прийняті з амбітною метою уніфікувати засоби впливу на життєвий цикл технологій ШІ з боку розробників та інших суб'єктів. В рамках процесу велика увага була сфокусована на проблематику генеративного ШІ, яка істотним чином визначатиме тенденції управління найближчим часом.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Басиста І.В., Удовенко Ж.В., Кулинич М.-М.А. Огляд тенденцій щодо штучного інтелекту та його перспективність для процесуальних рішень у перебігу кримінального провадження. *Науковий вісник Ужгородського Національного Університету. Серія: Право.* 2024. Випуск 81: частина 3. С. 19–38. <https://doi.org/10.24144/2307-3322.2024.81.3.3>.
2. Марущак А.І. Вплив міжнародних процесів регулювання штучного інтелекту на інформаційне право України. *Інформація і право.* 2023. № 4 (47). С. 55–63.
3. Подільчак О.М. Міжнародно-правове правове регулювання протидії злочинності в інформаційно-комунікаційних системах. *Наукові перспективи.* 2023. № 12 (42). С. 562–576. [https://doi.org/10.52058/2708-7530-2023-12\(42\)-562-576](https://doi.org/10.52058/2708-7530-2023-12(42)-562-576).
4. Bremmer I., Suleyman M. The AI Power Paradox. Can States Learn to Govern Artificial Intelligence – Before It's Too Late? *Foreign Affairs.* 2023. 102. Pp. 26–43.
5. Chan K. Multilateralism in the digital age: how to make progress on global digital governance. G7 & G20 Information Centre. Toronto, 2023. 2 p.
6. A. Ema et al. International Collaboration in AI Governance Key Considerations of the Council of Europe's AI Convention and Japan's Response. *IFI Policy Recommendation.* 2023. No. 25 (October). Pp. 1–13. [https://ifi.u-tokyo.ac.jp/en/wp-content/uploads/2023/10/policy\\_recommendation\\_tg\\_20231031e.pdf](https://ifi.u-tokyo.ac.jp/en/wp-content/uploads/2023/10/policy_recommendation_tg_20231031e.pdf) (дата звернення: 15.04.2024).
7. Japan pushes for basic AI rules at G-7 tech meeting. *The Japan Times.* URL: <https://www.japantimes.co.jp/news/2016/04/29/national/japan-pushes-basic-ai-rules-g-7-tech-meeting/> (дата звернення: 15.04.2024).
8. Ministerial Declaration – The G7 Digital and Tech Ministers' Meeting. *G7/G20 Documents Database.* URL: <https://g7g20-documents.org/database/document/2023-g7-japan-ministerial-meetings-ict-ministers-ministers-language-ministerial-declaration-the-g7-digital-and-tech-ministers-meeting#section-4> (дата звернення: 15.04.2024).
9. G7 Digital and Tech Track Annex 5. G7 Action Plan for promoting global interoperability between tools for trustworthy AI. *G7/G20 Documents Database.* URL: <https://g7g20-documents.org/database/document/2023-g7-japan-ministerial-meetings-ict-ministers-ministers-annex-g7-digital-and-tech-track-annex-5-g7-action-plan-for-promoting-global-interoperability-between-tools-for-trustworthy-ai> (дата звернення: 15.04.2024).
10. G7 Hiroshima Leaders' Communiqué. *G7/G20 Documents Database.* [https://g7g20-documents.org/fileadmin/G7G20\\_documents/2023/G7/Japan/Leaders/1%20Leaders'%20Language/G7%20Hiroshima%20Leaders%E2%80%99%20Communiqu%C3%A9\\_20052023.pdf](https://g7g20-documents.org/fileadmin/G7G20_documents/2023/G7/Japan/Leaders/1%20Leaders'%20Language/G7%20Hiroshima%20Leaders%E2%80%99%20Communiqu%C3%A9_20052023.pdf) (дата звернення: 15.04.2024).
11. G7 Hiroshima AI Process G7 Digital & Tech Ministers' Statement. *G7/G20 Documents Database.* URL: <https://g7g20-documents.org/database/document/2023-g7-japan-ministerial-meetings-ict-ministers-ministers-language-g7-hiroshima-ai-process-g7-digital-tech-ministers-statement> (дата звернення: 15.04.2024).
12. Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System. *G7/G20 Documents Database.* URL: [https://g7g20-documents.org/fileadmin/G7G20\\_documents/2023/G7/Japan/Leaders/2%20Leaders'%20Annex/Hiroshima%20Process%20International%20Guiding%20Principles%20for%20Organizations%20Developing%20Advanced%20AI%20System\\_30102023.pdf](https://g7g20-documents.org/fileadmin/G7G20_documents/2023/G7/Japan/Leaders/2%20Leaders'%20Annex/Hiroshima%20Process%20International%20Guiding%20Principles%20for%20Organizations%20Developing%20Advanced%20AI%20System_30102023.pdf) (дата звернення: 15.04.2024).
13. Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems. *G7/G20 Documents Database.* URL: [https://g7g20-documents.org/fileadmin/G7G20\\_documents/2023/G7/Japan/Leaders/2%20Leaders'%20Annex/Hiroshima%20Process%20International%20Code%20of%20Conduct%20for%20Organizations%20Developing%20Advanced%20AI%20](https://g7g20-documents.org/fileadmin/G7G20_documents/2023/G7/Japan/Leaders/2%20Leaders'%20Annex/Hiroshima%20Process%20International%20Code%20of%20Conduct%20for%20Organizations%20Developing%20Advanced%20AI%20)

- Systems\_30102023.pdf (дата звернення: 15.04.2024).
14. Hiroshima AI Process G7 Digital & Tech Ministers' Statement. *Ministry of Internal Affairs and Communications*. URL: [https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document02\\_en.pdf](https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document02_en.pdf) (дата звернення: 15.04.2024).
  15. G7 Leaders' Statement on the Hiroshima AI Process. *G7/G20 Documents Database*. URL: [https://g7g20-documents.org/fileadmin/G7G20\\_documents/2023/G7/Japan/Leaders/1%20Leaders'%20Language/G7%20Leaders%20Statement%20on%20the%20Hiroshima%20AI%20Process\\_30102023.pdf](https://g7g20-documents.org/fileadmin/G7G20_documents/2023/G7/Japan/Leaders/1%20Leaders'%20Language/G7%20Leaders%20Statement%20on%20the%20Hiroshima%20AI%20Process_30102023.pdf) (дата звернення: 15.04.2024).
  16. G7 Leaders' Statement. *The European Council*. URL: <https://www.consilium.europa.eu/media/68615/g7-leaders-statement-2023-12-6.pdf> (дата звернення: 15.04.2024).
  17. G7 Hiroshima Process on Generative Artificial Intelligence (AI). Towards a G7 Common Understanding on Generative AI. *OECD iLibrary*. URL: [https://www.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative-artificial-intelligence-ai\\_bf3c0c60-en](https://www.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en) (дата звернення: 15.04.2024).