DOI: 10.33498/louu-2025-07-122

Олег Посикалюк

кандидат юридичних наук, доцент,
завідувач сектору організації державної влади та місцевого самоврядування відділу з питань правової політики, організації публічної влади Дослідницької служби Верховної Ради України, перший заступник головного редактора юридичного журналу "Право України"
(Київ, Україна)
ORCID ID: https://orcid.org/0000-0002-8841-8481
oleg.posykaliuk@gmail.com

УДК 347.78

ШТУЧНИЙ ІНТЕЛЕКТ І ПЕРСОНАЛЬНІ ДАНІ: ДО ПИТАННЯ ПРО ПОШУК БАЛАНСУ ІНТЕРЕСІВ

Анотація. У статті досліджується фундаментальний конфлікт між потребами сучасних систем штучного інтелекту (ШІ), зокрема великих мовних моделей (LLM), у величезних масивах даних ("data hunger") та стандартами захисту персональних даних.

Актуальність дослідження зумовлена стрімким впровадженням генеративного ШІ (ChatGPT, Gemini, Copilot), яке створює системні ризики для приватності, включаючи алгоритмічну упередженість, деанонімізацію та непрозорість "чорних скриньок". Технологічні інновації значно випереджають розвиток адекватного правового регулювання.

Метою статті є дослідження правових, технічних та організаційних заходів, які вживаються для пошуку балансу між інтересами розробників ШІ та правами суб'єктів персональних даних.

У статті проаналізовано: 1) невідповідність ключових принципів Загального регламенту про захист даних (мінімізація даних, обмеження мети) архітектурі великих мовних моделей LLM та доповнюючу роль Акта ЄС про ШІ; 2) ефективність технічних заходів (PETs), таких як синтетичні дані, довірені середовища виконання та федеративне навчання; 3) організаційні підходи та політики приватності OpenAI, Google і Microsoft.

У висновках наголошується, що компанії впровадили дихотомію в гарантіях захисту: надійні договірні умови для корпоративних клієнтів і модель "opt-out" для масових споживачів, що перекладає тягар захисту приватності на самого користувача. Впровадження технічних заходів залишається частковим через їх ресурсомісткість. Констатується, що наявних заходів недостатньо для забезпечення реального балансу інтересів для індивідуальних користувачів.

Ключові слова: штучний інтелект; персональні дані; Загальний регламент про захист даних; Акт ЄС про штучний інтелект; великі мовні моделі; політика приватності; ChatGPT; Gemini; Copilot.

Штучний інтелект (далі – ШІ), з одного боку, обіцяє еру безпрецедентного технологічного прогресу, здатного трансформувати практично кожен аспект суспільного життя, від підвищення ефективності бізнесу до революційних проривів у медицині та науці, а з іншого – створює нові, складні та часто непередбачувані ризики для фундаментальних прав і свобод людини, насамперед

права на приватність. Пояснюється це тим, що для прогресу систем штучного інтелекту потрібні величезні масиви даних, значна частина яких є персональними та чутливими за своєю природою. Проблема полягає не в ШІ як такому, а в його залежності від даних: чим більше даних обробляє модель, тим ефективнішою вона стає, але водночас тим вищими стають ризики для приватності. Варто зауважити, що така залежність є нелінійною: кожне покращення моделі ШІ може підвищувати ступінь ризиків для конфіденційності, посилюючи причинно-наслідковий зв'язок між інтенсифікацією збору та обробки персональних даних та загрозами для особистої сфери.

Отже, виникає фундаментальний конфлікт між потребою сучасних систем ШІ в наборах даних і стандартами захисту персональних даних. Традиційні підходи до захисту даних, розроблені в епоху менш інтенсивної та всеохопної обробки інформації, виявляються недостатньо гнучкими та ефективними перед обличчям нових викликів, які ставить ШІ. Це створює прогалину, де технологічні інновації значно випереджають розвиток правового регулювання.

Потреба систем ШІ у величезних обсягах даних є не довільною вимогою, а фундаментальною технічною передумовою їхнього функціонування та розвитку. Особливо це стосується великих мовних моделей (LLM), які лежать в основі багатьох сучасних генеративних застосунків. Їхні можливості безпосередньо залежать від якості та масштабу даних, на яких вони навчаються. Для того, щоб системи ШІ могли адекватно розуміти та генерувати людську мову (Natural Language Understanding, NLU), їм необхідно навчатися на текстах, створених людьми, які є носіями не лише формальних правил граматики та синтаксису, а й незліченних семантичних відтінків, контекстуальних зв'язків, культурних нюансів та ідіоматичних виразів[1]. Персональні дані, такі як листування в соціальних мережах, відгуки на товари, форуми, блоги та інші форми онлайн-комунікації, є найбагатшим і найрізноманітнішим джерелом таких природних мовних патернів. Вони відображають мову в її "живому" вигляді, що дає змогу моделям навчатися не лише формальним структурам, а й прагматиці спілкування.

Ще одним із найскладніших завдань для ШІ є набуття так званого здорового глузду – здатності робити логічні, правдоподібні висновки про повсякденні ситуації та події, які для людини є очевидними. Ця здатність не може бути запрограмована у вигляді жорстких правил через нескінченну варіативність реального світу[2]. ШІ набуває її шляхом аналізу величезної кількості прикладів людської взаємодії, описів подій і знань про світ, які ча-

[1] Див. докладніше: A Rajasekharan, Y Zeng, P Padalkar, G Gupta, 'Reliable Natural Language Understanding with Large Language Models and Answer Set Programming' [2023] 385 Electronic Proceedings in Theoretical Computer Science 274–287.

[2] Див. докладніше: M Sap, V Shwartz, A Bosselut, Y Choi, D Roth, 'Commonsense Reasoning for Natural Language Processing', *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts* (Association for Computational Linguistics 2020) 27–33.

сто містяться саме в персональних даних. Без доступу до таких даних моделі схильні робити абсурдні помилки, що суттєво обмежує їхню надійність та практичне застосування в реальних умовах.

Крім того, сучасні дослідження LLM виявили феномен "емерджентних властивостей" – несподіваних нових можливостей, які з'являються у моделей лише після досягнення певного критичного обсягу навчальних даних. Такі складні навички, як арифметичне мислення, розуміння нюансів мови або здатність до логічних висновків, не проявляються в менших моделях, але раптово виникають у великих[3]. Це пояснює, чому розробники прагнуть використовувати дедалі більші датасети, які можуть містити сотні мільярдів слів і мати розмір у десятки терабайт. Масштабування є єдиним відомим на сьогодні способом досягти проривних результатів у складних завданнях, таких як медична діагностика, наукові дослідження чи створення складного програмного коду. Тож потреба в персональних даних для ШІ не є статичною; вона зростає разом з амбіціями щодо його можливостей. Перехід від простих завдань, як-от класифікація тексту, до складних, таких як ведення змістовного діалогу, вимагає не просто більше даних, а даних вищої якості. Чим більше ми прагнемо, щоб ШІ був схожим на людину у своїх когнітивних здібностях, тим більше він повинен "вбирати" людський досвід, зафіксований у персональних даних.

Водночас обробка персональних даних системами ШІ породжує багатогранний спектр ризиків. Ці ризики є системними, часто вбудованими в саму архітектуру та принципи роботи сучасних алгоритмів. Одним із найбільш соціально значущих ризиків є алгоритмічна упередженість. Системи ШІ, навчаючись на історичних даних, що відображають існуючі в суспільстві стереотипи та нерівності, не просто відтворюють їх, а й можуть посилювати, надаючи дискримінаційним практикам вигляду об'єктивності та технологічної нейтральності[4]. Характеристики ШІ посилюють традиційні ризики, пов'язані з обробкою персональних даних: надлишковий збір, нецільове використання, надмірне зберігання тощо. Крім того, ШІ значно підвищує ризик деанонімізації, тобто повторної ідентифікації осіб у наборах даних, які вважалися анонімними. Алгоритми можуть виявляти складні патерни та кореляції, дозволяючи поєднувати нібито знеособлені дані з іншими загальнодоступними джерелами для встановлення особистості. Значний ризик також створює здатність LLM робити висновки про чутливі персональні дані з нечутливого тексту, наданого під час взаємодії, тобто нову інформацію про особу, яку ШІ створює на основі аналізу наявних даних (наприклад, висновки про стан здоров'я, політичні погляди тощо). Більше того, багато сучасних моделей ШІ, особливо ті, що базуються на глибоких нейронних

[3] P Yu, H Xu, X Hu, C Deng, 'Leveraging Generative AI and Large Language Models: A Comprehensive Roadmap for Healthcare Integration' [2023] 11(20) Healthcare 2776.

[4] A Almufarreh, A Ahmad, M Arshad, O Choo Wou, R Elechi, 'Ethical implications of ChatGPT and other large language models in academia' [2025] 8 Frontiers in Artificial Intelligence doi: 10.3389/frai.2025.1615761.

мережах, функціонують як "чорні скриньки". Це означає, що навіть їхнім розробникам складно або неможливо точно пояснити, чому модель прийняла те чи інше конкретне рішення.

Окремі питання пошуку балансу між розвитком технологій штучного інтелекту та захистом персональних даних уже досліджували такі вітчизняні автори: М. Бєлова і Д. Бєлов[5], К. Резворович і М. Береда[6], О. Пунда і Д. Арзянцева[7], А. Колесніков, Я. Чапельський, В. Будник і Ю. Коженьовський[8], В. Базалицький[9], В. Некрутенко[10], С. Брайчевський[11] та ін. Деякі науковці присвячували свої праці особливостям обробки персональних даних при використанні чат-ботів зі штучним інтелектом на прикладі ChatGPT, зокрема О. Заярний і Ю. Деркаченко[12], А. Гачкевич[13]. У цій праці проведемо дослідження технічних, організаційних і правових заходів врегулювання конфлікту між інтересами розробників і операторів ШІ та правами суб'єктів персональних даних на прикладі провідних генеративних моделей штучного інтелекту ChatGPT, Gemini, Copilot. Цей вибір зумовлений їх місцем на ринку, широким використанням у світі та значним впливом на всю індустрію ШІ через належність найбільшим технологічним компаніям (Google, Microsoft, OpenAI).

Слід зазначити, що потреба ШІ в даних та способи їх використання зумовлюють необхідність враховувати вимоги Загального регламенту про захист даних (GDPR) – фундаментального елементу європейської (а отже, й глобальної[14]) нормативної системи управління даними. Варто звернути

[5] М Бєлова, Д Бєлов, 'Виклики та загрози захисту персональних даних у роботі зі штучним інтелектом' [2023] 79(2) Науковий вісник Ужгородського національного університету. Серія: Право 17–22.

[6] К Резворович, М Береда, 'Вплив штучного інтелекту на правову систему та захист персональних даних у цифрову епоху' [2024] 4(4) Успіхи і досягнення у науці. Серія "Право" 241–248.

[7] О Пунда, Д Арзянцева, 'Забезпечення захисту персональних даних фізичних осіб в умовах розвитку штучного інтелекту' [2024] 2(30) Наука і техніка сьогодні. Серія "Право" 132–142.

[8] А Колесніков, Я Чапельський, В Будник, Ю Коженьовський, 'Трансформація системи захисту персональних даних під впливом розвитку технологій штучного інтелекту' [2025] 2 Економіка. Фінанси. Право 106–109.

[9] В Базалицький, 'Врегулювання питання обробки персональних даних штучним інтелектом у Загальному регламенті із захисту персональних даних (GDPR)' [2024] 6(24) Актуальні питання у сучасній науці. Серія "Право" 406–419.

[10] В Некрутенко, 'До питання систематизації ризиків, спричинених обробленням персональних даних із використанням технології штучного інтелекту' [2021] 4(119) Вісник Київського національного університету імені Тараса Шевченка (юридичні науки) 53–58.

[11] С Брайчевський, 'Проблема персональних даних в системах Інтернету речей з елементами штучного інтелекту' [2019] 3 Інформація і право 61–67.

[12] О Заярний, Ю Деркаченко, 'Деякі особливості обробки персональних даних при використанні чат-ботів зі штучним інтелектом на прикладі ChatGPT' [2023] 29 Юридичний бюлетень 55–62.

[13] А Гачкевич, 'Нагляд національних органів ЄС із захисту даних за обробкою персональних даних системами штучного інтелекту (на прикладі ChatGPT)' [2025] 2(4) Аналітично-порівняльне правознавство 154–160.

[14] Це пояснюється, з одного боку, екстериторіальною дією Загального регламенту про захист даних, яка прямо закріплена в статті 3. З іншого – тим, що передавання персональних даних до третіх країн або міжнародних організацій регулюється главою V Загального регламенту про захист даних та повинно відповідати загальному принципу такого передання: 'Будь-яка передача персональних даних, що обробляються або призначені для обробки після передачі до третьої країни або міжнародної організації, може здійснюватися лише за умови дотримання інших положень цього Регламенту та виконання умов, встановлених у цьому розділі, як контролером, так і оператором, включно з подальшими передачами персональних даних з третьої країни або міжнародної організації до іншої третьої країни

увагу на вимогу ст. 6 GDPR, згідно з якою будь-яка обробка персональних даних має ґрунтуватися на одній з шести правових умов. Власне необхідність забезпечення досліджуваного нами балансу випливає з пункт "f" частини першої цієї статті, відповідно до якого обробка персональних даних є законною, якщо є необхідною для реалізації законних інтересів, що переслідуються контролером або третьою стороною, за винятком випадків, коли такі інтереси переважаються інтересами або основоположними правами й свободами суб'єкта персональних даних. Пункт 47 преамбули пояснює, що наявність такого законного інтересу потребуватиме ретельної оцінки, включно з тим, чи може суб'єкт даних обґрунтовано очікувати в момент та в контексті збору персональних даних, що обробка з цією метою може відбутися. При цьому суб'єкт персональних даних повинен мати право заперечувати проти обробки будь-яких персональних даних. Контролер повинен продемонструвати, що його законний інтерес має перевагу над інтересами або основоположними правами і свободами суб'єкта персональних даних (п. 69 преамбули).

Водночас виклики, які постають перед регулюванням ШІ, є значно складнішими, ширшими та менш визначеними, ніж ті, що є предметом регулювання GDPR. У відповідь на це було прийнято Акт ЄС про ШІ. Цей Акт не замінює GDPR, а доповнює його, створюючи специфічні правила для розробки та використання систем ШІ. Необхідність забезпечувати баланс між розвитком технологій ШІ та захистом персональних даних випливає зі ст. 1 Акта ЄС про ШІ, в якій закріплено мету правового регулювання щодо ШІ: 'сприяння впровадженню людиноцентричного та надійного штучного інтелекту (ШІ), водночас забезпечуючи високий рівень захисту здоров'я, безпеки, основоположних прав'[15]. Положення преамбули цього регламенту допомагають конкретизувати цю мету. З одного боку, встановлення гармонізованих правил щодо ШІ є необхідним для сприяння розробці, використанню та впровадженню ШІ на внутрішньому ринку (п. 8):

> Ці правила мають бути чіткими та надійними у захисті основоположних прав, підтримувати нові інноваційні рішення, створювати умови для європейської екосистеми державних і приватних суб'єктів, які розробляють системи ШІ відповідно до

або іншої міжнародної організації. Усі положення цього розділу застосовуються з метою забезпечення того, щоб рівень захисту фізичних осіб, гарантований цим Регламентом, не був знижений'. Див.: Регламент Європейського Парламенту і Ради (ЄС) 2016/679 від 27 квітня 2016 р. про захист фізичних осіб у зв'язку з опрацюванням персональних даних і про вільний рух таких даних, та про скасування Директиви 95/46/ЄС (Загальний регламент про захист даних) <https://zakon.rada.gov.ua/laws/show/984_008-16#Text> (дата звернення 27.07.2025).

[15] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> (accessed 27.07.2025).

цінностей Союзу, та розкривати потенціал цифрової трансформації в усіх регіонах Союзу[16].

З іншого, акцентується увага, що суб'єкти персональних даних продовжують користуватися всіма наданими їм правами та гарантіями (п. 10):

Гармонізовані правила щодо розміщення на ринку, введення в експлуатацію та використання систем ШІ, встановлені цим Регламентом, повинні сприяти ефективному застосуванню та уможливлювати реалізацію прав суб'єктів даних та інших засобів правового захисту, гарантованих законодавством Союзу про захист персональних даних та інших основоположних прав[17].

Акт ЄС про ШІ та GDPR зосереджуються на дуже різних підходах до регулювання: GDPR застосовує підхід, що ґрунтується на правах, формулюючи конкретні права, які особи мають щодо кожної транзакції персональними даними; Акт ЄС про ШІ будує свої правила навколо підходу, що ґрунтується на ризиках, який накладає вимоги на власників та операторів систем ШІ відповідно до рівня ризиків, які ці системи створюють. Більше того, підхід, що ґрунтується на правах, відображений у GDPR, регулює персональні дані як вхідні дані для обробки, тоді як підхід, що ґрунтується на ризиках, відображений у Акті ЄС про ШІ, зосереджується на використанні продуктів ШІ як вихідних даних[18].

Варто розуміти, що з точки зору захисту персональних даних важливими є усі життєві цикли моделі ШІ, які охоплюють принаймні три етапи. Перший етап – це збір великої кількості даних як навчальних даних. Ці дані можуть містити персоніфіковану або неперсоніфіковану інформацію. У певних випадках цей процес використовує надзвичайно великі набори даних, що робить складним, якщо не неможливим, розрізнення між різними категоріями даних. Наприклад, ChatGPT було розроблено з використанням величезних обсягів даних, що є у вільному доступі в Інтернеті. Другий етап – це власне навчання моделі з використанням зібраних даних, результатом чого є сконфігурована модель. Третій етап – це застосування моделі, що означає, що навчена модель застосовується до конкретних випадків або осіб, перетворюючи модель на інструмент, який обчислює конкретний вихідний результат у відповідь на вхідні дані[19]. Кожен із зазначених етапів створює значні ризики для захисту персональних даних.

По-перше, генеративні моделі ШІ зазвичай навчаються на мільярдах, якщо не на сотнях мільярдів, параметрів і вимагають великих обсягів навчальних даних. Натомість захист персональних даних ґрунтується на ідеї

[16] Artificial Intelligence Act (n 15).
[17] Ibid.
[18] J Wolff, W Lehr, C S Yoo, 'Lessons from GDPR for AI Policymaking' [2024] 27(4) Virginia Journal of Law & Technology 20, 22.
[19] R Mühlhoff, H Ruschemeier, 'Regulating AI with Purpose Limitation for Models' [2024] 1 Journal of AI Law and Regulation 24–39.

індивідуального контролю, дозволяючи суб'єктам персональних даних керувати власною персональною інформацією. Але моделі ШІ, навчені на безпрецедентно великих наборах даних, унеможливлюють ручну ідентифікацію або навіть перевірку відповідності оброблених даних вимогам GDPR, таким чином, містять потенціал для порушень конфіденційності та захисту даних[20]. Крім того, принцип мінімізації даних, викладений в п. "c" ч. 1 ст. 5 GDPR, який вимагає обробляти лише ті дані, які є абсолютно необхідними для конкретної мети, прямо суперечить архітектурі LLM. Ці моделі потребують масивних, надлишкових наборів даних для ефективного навчання та досягнення високої точності (так званий спрага даних, англ. – "data hunger"[21]).

По-друге, однією з найбільших проблем у сфері конфлікту між ШІ та захистом персональних даних є відсутність прозорості багатьох систем ШІ, які часто називають "чорними скриньками". Зокрема, зі складними моделями машинного навчання часто важко зрозуміти, як саме приймаються рішення або отримуються результати[22]. Ця відсутність простежуваності може значно ускладнити виконання зобов'язань щодо прозорості, передбачених GDPR (ст. 12–14), та перешкоджати ефективному здійсненню прав суб'єктів даних, таких як право доступу (ст. 15 GDPR) або право на пояснення автоматизованих рішень (ст. 22 GDPR).

По-третє, великі мовні моделі охоплюють широкий спектр цілей, застосувань та операційних середовищ. Згідно з п. 63 ч. 1 ст. 3 Акта ЄС про ШІ однією з особливостей моделей ШІ загального визнається їх здатність компетентно виконувати широкий спектр окремих завдань і можливість бути інтегрованою в різноманітні наступні системи або програми. Зокрема, коли моделі ШІ стають доступними для численних третіх сторін через інтерфейс, забезпечення того, що цілі, для яких використовується ця модель, а отже, і її дані, є сумісними з початковими цілями збору даних, стає складним, якщо не неможливим[23]. Цей аспект важко узгодити з принципом обмеження мети GDPR (ч. 4 ст. 6), оскільки у сфері ШІ поширеною є практика використання старих наборів даних для навчання нових моделей або для цілей, які неможливо було передбачити на момент збору даних (перепрофілювання даних, англ. – "data repurposing"[24]).

Це спонукає розробників та операторів систем ШІ до застосування різних технічних та організаційних заходів, спрямованих на мінімізацію ри-

[20] H Ruschemeier, 'Generative AI and data protection' [2025] 1 Cambridge Forum on AI: Law and Governance doi:10.1017/cfl.2024.2.

[21] T Davies, 'Data Hunger: The Deep Connection Between the AI Chatbot and the Human' [2025] 44(1) IEEE Technology and Society Magazine 43–50.

[22] C Schwabe, 'AI and data protection in practice – between innovation and regulation' (7 April 2025, Robin Data GmbH) <https://www.robin-data.io/en/data-protection-and-data-security-academy/wiki/ki-und-datenschutz-praxisleitfaden> (accessed 27.07.2025).

[23] Mühlhoff, Ruschemeier (n 19) 24–39.

[24] J Parsons, R Lukyanenko, B Greenwood, C Cooper, 'Understanding and Improving Data Repurposing' [2025] MIS Quarterly 1–50 https://doi.org/10.48550/arXiv.2506.09073.

зиків порушення прав суб'єктів персональних даних. На необхідності застосування технічних заходів для гарантування захисту персональних даних вказує також Акт ЄС про ШІ. Зокрема, п. 69 преамбули вказує, що для забезпечення відповідності принципам мінімізації даних і захисту даних за задумом і за замовчуванням слід вживати заходів, які 'можуть включати не лише анонімізацію та шифрування, але й використання технологій, які дають змогу переносити алгоритми до даних і навчати системи ШІ без передачі між сторонами або копіювання самих необроблених чи структурованих даних, без шкоди для вимог щодо управління даними'[25].

У доповіді Організації економічного співробітництва та розвитку зазначається, що основна роль технологій підвищення конфіденційності ("privacy-enhancing technologies", або PETs), які використовуються в контексті штучного інтелекту, – це забезпечення довіри під час спільної розробки та обміну моделями ШІ[26]. Ці технології охоплюють весь процес – від збору даних і навчання моделей до їхнього розгортання та використання. Виділяють такі ключові PETs:

– Синтетичні дані (Synthetic Data – SD): створення штучних наборів даних, що імітують статистичні властивості реальних даних. Це дає змогу оптимізувати та тестувати моделі без використання справжніх, потенційно чутливих даних.

– Диференційна приватність (Differential Privacy – DP): додавання математичного "шуму" до даних для зменшення ризику повторної ідентифікації окремих осіб, зберігаючи при цьому загальну корисність даних. Часто використовується в поєднанні з синтетичними даними.

– Гомоморфне шифрування (Homomorphic Encryption – HE): дає змогу проводити обчислення безпосередньо над зашифрованими даними без необхідності їх розшифровувати. Це захищає конфіденційність даних під час їх обробки.

– Багатосторонні обчислення (Multi-Party Computation – MPC): дають змогу кільком сторонам спільно аналізувати свої дані, не розкриваючи їх одна одній. Кожна сторона дізнається лише кінцевий результат обчислень.

– Довірені середовища виконання (Trusted Execution Environments – TEEs): створюють захищені апаратні анклави на процесорі, де дані та код можуть оброблятися ізольовано від основної операційної системи, захищаючи їх навіть від хмарного провайдера.

– Федеративне навчання (Federated Learning – FL): метод навчання глобальної моделі ШІ шляхом агрегації оновлень від локальних моделей, які тренуються на пристроях користувачів (наприклад, телефонах). При цьому вихідні "сирі" дані ніколи не залишають пристрій[27].

[25] Artificial Intelligence Act (n 15).
[26] OECD (2025), "Sharing trustworthy AI models with privacy-enhancing technologies", OECD Artificial Intelligence Papers, No. 38, OECD Publishing, Paris, https://doi.org/10.1787/a266160b-en.
[27] OECD (2025), "Sharing trustworthy AI models with privacy-enhancing technologies" (n 26).

Слід мати на увазі, що жоден із таких технічних заходів не є універсальним, часто їх потрібно комбінувати, щоб компенсувати недоліки окремих технологій і знайти баланс між корисністю даних, ефективністю та рівнем захисту приватності. Більше того, практичне впровадження багатьох PETs у моделі масштабу LLM стикається з серйозними перешкодами, оскільки деякі з них є надзвичайно ресурсомісткими (гомоморфне шифрування та багатосторонні обчислення) або можуть призводити до зниження точності та загальної якості моделі ШІ (диференційна приватність).

Це підтверджується також і тим, що у досліджуваних нами LLM впровадження PETs наразі далеке від масштабного. Так, OpenAI є єдиною компанією, яка відкрито підтверджує використання синтетичних даних як ключового компонента для навчання та налаштування своїх передових моделей. Зокрема, у технічній документації до DALL-E 3, де "синтетичні підписи" до зображень використовувалися для покращення слідування інструкціям[28]. Модель Phi-3 від Microsoft малої мови програмування (SLM) є прикладом того, як синтетичні дані можуть сприяти відповідальній розробці штучного інтелекту, даючи змогу створювати потужні мовні моделі без шкоди для конфіденційності. Phi-3 використовує комбінацію веб-даних "підручникової якості" та синтетичного контенту, згенерованого LLM, створюючи стратегічний підхід, який не потребує реальних персональних даних[29]. Microsoft також анонсувала сервіс "конфіденційного інференсу" для Azure AI, що використовує TEEs для наскрізного захисту запитів до моделей ШІ, включно з тими, що лежать в основі Copilot[30]. Google також активно інтегрує технології конфіденційних обчислень у свою екосистему, і це стосується безпосередньо моделей Gemini. Google Cloud пропонує широкий портфель продуктів Confidential Computing, включаючи Confidential VMs та Confidential Space – захищене середовище для спільної роботи з даними, де жодна зі сторін (включаючи Google) не має доступу до даних іншої[31].

Акт ЄС про ШІ також акцентує увагу важливості впровадження організаційних заходів, спрямованих на доступ до високоякісних даних, які відіграють життєво важливу роль у забезпеченні структури та продуктивності багатьох систем ШІ, особливо коли використовуються методи, що передбачають навчання моделей.

[28] Improving Image Generation with Better Captions (OpenAI) <https://cdn.openai.com/papers/dall-e-3.pdf> (accessed 27.07.2025).

[29] G Afonja, R Sim, Z Lin, H A Inan, S Yekhanin, 'The Crossroads of Innovation and Privacy: Private Synthetic Data for Generative AI' (May 29, 2024) <https://www.microsoft.com/en-us/research/blog/the-crossroads-of-innovation-and-privacy-private-synthetic-data-for-generative-ai> (accessed 27.07.2025).

[30] T Numoto, 'Microsoft Trustworthy AI: Unlocking human potential starts with trust' (Sep 24, 2024) <https://blogs.microsoft.com/blog/2024/09/24/microsoft-trustworthy-ai-unlocking-human-potential-starts-with-trust> (accessed 27.07.2025).

[31] Confidential Computing <https://cloud.google.com/security/products/confidential-computing?hl=uk> (accessed 27.07.2025).

> Високоякісні набори даних для навчання, валідації та тестування вимагають впровадження належних практик управління та менеджменту даних <…> З метою полегшення дотримання законодавства Союзу про захист даних, такого як Регламент (ЄС) 2016/679, практики управління та менеджменту даних повинні включати, у випадку персональних даних, прозорість щодо первинної мети збору даних <…> Вимоги, пов'язані з управлінням даними, можуть бути виконані шляхом звернення до третіх сторін, які пропонують сертифіковані послуги з дотримання вимог, включаючи перевірку управління даними, цілісності наборів даних, а також практик навчання, валідації та тестування даних, за умови забезпечення відповідності вимогам цього Регламенту щодо даних[32].

Технології самі по собі є недостатніми без надійних організаційних та управлінських структур. Організаційні заходи включають: розроблення політик та процедур захисту персональних даних; проведення регулярних аудитів безпеки; навчання персоналу безпечному обробленню персональних даних; впровадження процедур реагування на інциденти безпеки[33] тощо. До прикладу, "Модельна рамкова програма захисту персональних даних у сфері штучного інтелекту", підготовлена Офісом Комісара з питань конфіденційності персональних даних Гонконгу[34], містить набір рекомендацій і найкращих практик щодо управління ШІ для захисту конфіденційності персональних даних для організацій, які закуповують, впроваджують і використовують будь-які типи систем ШІ. Модельна рамкова програма містить рекомендовані заходи, які охоплюють весь життєвий цикл ШІ та групуються в таких чотирьох напрямах: визначення стратегії та управління у сфері ШІ (розробка стратегії ШІ, створення органів з управління ШІ, навчання та підвищення кваліфікації всіх співробітників); оцінка ризиків та людський нагляд (аналіз потенційного впливу рішень ШІ на права суб'єкта персональних даних і вжиття заходів з їх мінімізації, визначення моделі людського нагляду, яка має бути пропорційною до рівня ризику); управління системою ШІ (належне документування, тестування моделей ШІ в процесі їх впровадження, безперервний моніторинг, внутрішній аудит); комунікація та взаємодія із зацікавленими сторонами (належне інформування суб'єктів персональних даних, функціонування каналів зворотного зв'язку). Важливим аспектом є впровадження принципу "приватність за дизайном", коли захист персональних даних враховується вже на етапі проєктування системи[35].

Підхід кожної досліджуваної компанії до управління ШІ глибоко інтегрований у її корпоративну стратегію. Так, задекларована місія OpenAI має забезпечити, щоб загальний ШІ приносив користь всьому людству. Під-

[32] Artificial Intelligence Act (n 15).
[33] Колесніков, Чапельський, Будник, Коженьовський (н 8) 107.
[34] Artificial Intelligence – Model Personal Data Protection Framework. Office of the Privacy Commissioner for Personal Data, Hong Kong 2024 <https://www.pcpd.org.hk/english/resources_centre/publications/files/ai_protection_framework.pdf> (accessed 27.07.2025).
[35] Колесніков, Чапельський, Будник, Коженьовський (н 8) 107.

порядкування неприбутковій організації й унікальна модель обмеженого прибутку означає, що зі зростанням потужності ШІ компанія може перерозподіляти прибутки від діяльності, щоб максимально збільшити соціальні та економічні вигоди від технологій штучного інтелекту[36]. Google обґрунтовує розробку своїх ШІ-продуктів набором публічно заявлених принципів ШІ: сміливі інновації – сприяння економічному прогресу, просування меж досліджень у сфері ШІ, використання ШІ для прискорення наукових відкриттів; відповідальна розробка та впровадження – запровадження належного людського контролю, розробка технічних рішень для зменшення ризиків, сприяння захисту приватності та безпеки; спільний прогрес – співпраця з дослідниками з індустрії та академічного середовища, взаємодія з урядами та громадянським суспільством[37]. Стандарт Microsoft щодо відповідального ШІ включає шість ключових принципів: справедливість, надійність і безпека, конфіденційність і захист, інклюзивність, прозорість та підзвітність[38].

У таких питаннях, як політика збору, використання і зберігання даних, а також механізми людського нагляду, усі досліджувані компанії демонструють дихотомію в гарантіях захисту даних між сервісами, орієнтованими на споживачів (безкоштовними або недорогими), та рішеннями корпоративного рівня. Корпоративні пропозиції надають надійні договірні гарантії, ізоляцію даних та адміністративні засоби контролю, які переважно відсутні в їхніх споживчих аналогах. При цьому спостерігається зміщення функції контролю за конфіденційністю даних з індивідуального кінцевого користувача в споживчих продуктах на адміністратора організації в корпоративних продуктах. У споживчому контексті конфіденційність є індивідуальним вибором. Юридичне зобов'язання платформи стосується окремого суб'єкта даних. Тому засоби контролю орієнтовані на користувача. У корпоративному контексті організація є контролером даних, які обробляють її співробітники. Юридичне зобов'язання платформи (як обробника даних) стосується організації. Тому засоби контролю повинні бути орієнтовані на адміністратора, щоб дати змогу організації застосовувати свої корпоративні політики. Організація не може покладатися на інструктування своїх співробітників щодо індивідуального управління налаштуваннями конфіденційності, натомість потребує надійних адміністративних засобів контролю[39].

Ключовою відмінністю досліджуваних платформ є те, що обсяг збору персональних даних великою мовною моделлю прямо пропорційний її інте-

[36] OpenAI Privacy Portal (January 12, 2024) <https://privacy.openai.com/policies> (accessed 27.07.2025).
[37] AI Principles <https://ai.google/principles> (accessed 27.07.2025).
[38] Microsoft Responsible AI: Principles and approach <https://www.microsoft.com/en-us/ai/principles-and-approach#responsible-ai-standard> (accessed 27.07.2025).
[39] Див.: Business data privacy, security, and compliance <https://openai.com/business-data> (accessed 27.07.2025); Generative AI in Google Workspace Privacy Hub <https://support.google.com/a/answer/15706919?hl=en> (accessed 27.07.2025); Enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat <https://learn.microsoft.com/en-us/copilot/microsoft-365/enterprise-data-protection> (accessed 27.07.2025).

грації в ширшу технологічну екосистему. Самостійні сервіси, як-от ChatGPT, мають вужчий обсяг збору даних порівняно з Gemini та Copilot, які використовують величезні масиви даних користувачів з допоміжних сервісів для покращення функціональності. Логіка цього явища полягає в тому, що архітектура Gemini та Copilot заснована на їхній унікальній перевазі – доступі до даних з Gmail, Docs, Android (у випадку Google) або Microsoft Graph (у випадку Microsoft). Ризик для приватності користувача більше не обмежується вікном чату, а поширюється на весь його цифровий слід у межах цієї екосистеми.

Політики приватності усіх трьох компаній значну увагу приділяють правовому режиму персональних даних і серед іншого передбачають такі ключові елементи.

– Використання персональних даних для навчання моделей ШІ. Для всіх трьох платформ спільним є те, що для індивідуальних, некорпоративних користувачів чати за замовчуванням використовуються для навчання та вдосконалення моделей ШІ. Ця модель "opt-out" перекладає тягар захисту приватності на користувача, який повинен знати про це налаштування та проактивно його вимкнути. Copilot пропонує найбільш деталізовані налаштування окремо для "Персоналізації" та "Тренування моделей". Це дає змогу користувачеві відмовитися від тренування моделей, але при цьому дозволити Copilot використовувати свою нещодавню історію розмов для надання персоналізованого досвіду[40].

– Анонімізація та деіндентифікація персональних даних. Поширеною практикою в усіх трьох політиках є використання юридично широко сформульованих зобов'язань щодо анонімізації без надання конкретних технічних деталей про використовувані методи. Так, політика приватності OpenAI передбачає знеособлення персональної інформації таким чином, щоб її більше не можна було використати для ідентифікації користувача, і використання такої інформації для аналізу ефективності їхніх послуг, удосконалення та додавання нових функцій, проведення досліджень та інших подібних цілей[41].

– Права суб'єктів персональних даних та їх реалізація. Усі три компанії визнають стандартні права суб'єктів даних, що випливають з таких регуляцій, як GDPR та Закону Каліфорнії про приватність споживачів (2018)[42], включаючи право на доступ, виправлення, перенесення та видалення персональних даних. При цьому може міститися застереження, що через особливість роботи LLM, реалізація деяких прав може бути ускладнена. Зокрема, в політиці приватності OpenAI міститься примітка, відповідно до якої

[40] Microsoft Copilot privacy controls <https://support.microsoft.com/en-us/topic/microsoft-copilot-privacy-controls-8e479f27-6eb6-48c5-8d6a-c134062e2be6> (accessed 27.07.2025).
[41] Privacy policy (27 June 2025) <https://openai.com/en-GB/policies/row-privacy-policy> (accessed 27.07.2025).
[42] California Consumer Privacy Act of 2018 <https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5> (accessed 27.07.2025).

‘З огляду на технічну складність роботи наших моделей, ми можемо не мати змоги виправити неточність у кожному випадку. У такому разі ви можете надіслати запит на видалення вашої персональної інформації з результатів ChatGPT, заповнивши відповідну форму’[43].

Проведене дослідження свідчить, що на сьогодні пошук балансу інтересів між потребами великих мовних моделей у великій кількості якісних персональних даних і вимогами їх захисту здійснюється принаймні на трьох рівнях:

– нормативний – зі змісту правових актів, таких як GDPR, випливає необхідність дотримання пропорційності між законними інтересами, що переслідуються розробниками та операторами ШІ, та інтересами або основоположними правами й свободами суб’єкта персональних даних, однак критерії такого балансу ще потребують свого вироблення, передусім на рівні судової практики;

– саморегульований – такі компанії, як Google, Microsoft, OpenAI, пропонують своїм клієнтам два підходи до захисту персональних даних: для масових споживачів правила збору, обробки, використання для навчання LLM, зберігання, видалення персональних даних регламентуються відповідними політиками приватності, на які споживачі погоджуються шляхом приєднання і які можуть бути змінені компаніями в односторонньому порядку із повідомленням споживачів; для корпоративних клієнтів пропонується значно вищий рівень контролю над збором, обробкою, використанням для навчання LLM, зберіганням, видаленням персональних даних, що може забезпечуватись положеннями двосторонніх договорів;

– організаційно-технічний – такі компанії, як Google, Microsoft, OpenAI, застосовують деякі організаційні і технічні заходи, спрямовані на забезпечення відповідності їх практик збору, обробки, використання для навчання LLM, зберігання, видалення персональних даних вимогам відповідних нормативних актів, однак впровадження таких заходів залишається частковим, оскільки окремі з них є або надто ресурсомісткими, або призводять до зниження точності моделей ШІ.

Конфлікт між інноваціями ШІ та персональними даними не зникне, а його балансування й надалі відбуватиметься за трьома паралельними напрямами: посилення регуляторного тиску шляхом прийняття національних законів про ШІ; персональні дані масового споживача залишатимуться предметом обміну на доступ до безкоштовних послуг, тоді як корпоративні клієнти отримуватимуть більш надійні та технічно верифіковані гарантії захисту персональних даних; фокус з кількості даних, необхідних для навчання загальних LLM, зміститься на якість та безпеку таких даних для розробки предметно-орієнтованих моделей ШІ.

[43] Privacy policy (n 41).

REFERENCES

Bibliography

Journal articles

1. Almufarreh A, Ahmad A, Arshad M, Choo Wou O, Elechi R, 'Ethical implications of ChatGPT and other large language models in academia' [2025] 8 Frontiers in Artificial Intelligence doi: 10.3389/frai.2025.1615761.
2. Davies T, 'Data Hunger: The Deep Connection Between the AI Chatbot and the Human' [2025] 44(1) IEEE Technology and Society Magazine 43–50.
3. Mühlhoff R, Ruschemeier H, 'Regulating AI with Purpose Limitation for Models' [2024] 1 Journal of AI Law and Regulation 24–39.
4. Parsons J, Lukyanenko R, Greenwood B, Cooper C, 'Understanding and Improving Data Repurposing' [2025] MIS Quarterly 1–50 https://doi.org/10.48550/arXiv.2506.09073.
5. Rajasekharan A, Zeng Y, Padalkar P, Gupta G, 'Reliable Natural Language Understanding with Large Language Models and Answer Set Programming' [2023] 385 Electronic Proceedings in Theoretical Computer Science 274–287.
6. Ruschemeier H, 'Generative AI and data protection' [2025] 1 Cambridge Forum on AI: Law and Governance doi:10.1017/cfl.2024.2.
7. Wolff J, Lehr W, Yoo C S, 'Lessons from GDPR for AI Policymaking' [2024] 27(4) Virginia Journal of Law & Technology 20, 22.
8. Yu P, Xu H, Hu X, Deng C, 'Leveraging Generative AI and Large Language Models: A Comprehensive Roadmap for Healthcare Integration' [2023] 11(20) Healthcare 2776.
9. Bazalytskyi V, 'Vrehuliuvannia pytannia obrobky personalnykh danykh shtuchnym intelektom u Zahalnomu rehlamenti iz zakhystu personalnykh danykh (GDPR)' [2024] 6(24) Aktualni pytannia u suchasnii nautsi. Ceriia "Pravo" 406–419 (in Ukrainian).
10. Bielova M, Bielov D, 'Vyklyky ta zahrozy zakhystu personalnykh danykh u roboti zi shtuchnym intelektom' [2023] 79(2) Naukovyi visnyk Uzhhorodskoho natsionalnoho universytetu. Seriia: Pravo 17–22 (in Ukrainian).
11. Braichevskyi S, 'Problema personalnykh danykh v systemakh Internetu rechei z elementamy shtuchnoho intelektu' [2019] 3 Informatsiia i pravo 61–67 (in Ukrainian).
12. Hachkevych A, 'Nahliad natsionalnykh orhaniv YeS iz zakhystu danykh za obrobkoiu personalnykh danykh systemamy shtuchnoho intelektu (na prykladi ChatGPT)' [2025] 2(4) Analitychno-porivnialne pravoznavstvo 154–160 (in Ukrainian).
13. Kolesnikov A, Chapelskyi Ya, Budnyk V, Kozhenovskyi Yu, 'Transformatsiia systemy zakhystu personalnykh danykh pid vplyvom rozvytku tekhnolohii shtuchnoho intelektu' [2025] 2 Ekonomika. Finansy. Pravo 106–109 (in Ukrainian).
14. Nekrutenko V, 'Do pytannia systematyzatsii ryzykiv, sprychynenykh obroblenniam personalnykh danykh iz vykorystanniam tekhnolohii shtuchnoho intelektu' [2021] 4(119) Visnyk Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka (iurydychni nauky) 53–58 (in Ukrainian).
15. Punda O, Arziantseva D, 'Zabezpechennia zakhystu personalnykh danykh fizychnykh osib v umovakh rozvytku shtuchnoho intelektu' [2024] 2(30) Nauka i tekhnika sohodni. Seriia "Pravo" 132–142 (in Ukrainian).
16. Rezvorovych K, Bereda M, 'Vplyv shtuchnoho intelektu na pravovu systemu ta zakhyst personalnykh danykh u tsyfrovu epokhu' [2024] 4(4) Uspikhy i dosiahnennia u nautsi. Seriia "Pravo" 241–248 (in Ukrainian).
17. Zaiarnyi O, Derkachenko Yu, 'Deiaki osoblyvosti obrobky personalnykh danykh pry vykorystanni chat-botiv zi shtuchnym intelektom na prykladi ChatGPT' [2023] 29 Yurydychnyi biuleten 55–62 (in Ukrainian).

Conference papers

18. Sap M, Shwartz V, Bosselut A, Choi Y, Roth D, 'Commonsense Reasoning for Natural Language Processing', Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts (Association for Computational Linguistics 2020) 27–33.

Websites

19. Afonja G, Sim R, Lin Z, Inan H A, Yekhanin S, 'The Crossroads of Innovation and Privacy: Private Synthetic Data for Generative AI' (May 29, 2024) <https://www.microsoft.com/en-us/research/blog/the-crossroads-of-innovation-and-privacy-private-synthetic-data-for-generative-ai> (accessed 27.07.2025).
20. AI Principles <https://ai.google/principles> (accessed 27.07.2025).
21. Artificial Intelligence – Model Personal Data Protection Framework. Office of the Privacy Commissioner for Personal Data, Hong Kong 2024 <https://www.pcpd.org.hk/english/resources_centre/publications/files/ai_protection_framework.pdf> (accessed 27.07.2025).
22. Business data privacy, security, and compliance <https://openai.com/business-data> (accessed 27.07.2025).
23. Confidential Computing <https://cloud.google.com/security/products/confidential-computing?hl=uk> (accessed 27.07.2025).
24. Generative AI in Google Workspace Privacy Hub <https://support.google.com/a/answer/15706919?hl=en> (accessed 27.07.2025); Enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat <https://learn.microsoft.com/en-us/copilot/microsoft-365/enterprise-data-protection> (accessed 27.07.2025).
25. Improving Image Generation with Better Captions (OpenAI) <https://cdn.openai.com/papers/dall-e-3.pdf> (accessed 27.07.2025).
26. Microsoft Copilot privacy controls <https://support.microsoft.com/en-us/topic/microsoft-copilot-privacy-controls-8e479f27-6eb6-48c5-8d6a-c134062e2be6> (accessed 27.07.2025).
27. Microsoft Responsible AI: Principles and approach <https://www.microsoft.com/en-us/ai/principles-and-approach#responsible-ai-standard> (accessed 27.07.2025).
28. Numoto T, 'Microsoft Trustworthy AI: Unlocking human potential starts with trust' (Sep 24, 2024) <https://blogs.microsoft.com/blog/2024/09/24/microsoft-trustworthy-ai-unlocking-human-potential-starts-with-trust> (accessed 27.07.2025).
29. OECD (2025), "Sharing trustworthy AI models with privacy-enhancing technologies", OECD Artificial Intelligence Papers, No. 38, OECD Publishing, Paris, https://doi.org/10.1787/a266160b-en.
30. OpenAI Privacy Portal (January 12, 2024) <https://privacy.openai.com/policies> (accessed 27.07.2025).
31. Privacy policy (27 June 2025) <https://openai.com/en-GB/policies/row-privacy-policy> (accessed 27.07.2025).
32. Schwabe C, 'AI and data protection in practice - between innovation and regulation' (7 April 2025, Robin Data GmbH) <https://www.robin-data.io/en/data-protection-and-data-security-academy/wiki/ki-und-datenschutz-praxisleitfaden> (accessed 27.07.2025).

Oleh Posykaliuk

ARTIFICIAL INTELLIGENCE AND PERSONAL DATA: ON THE SEARCH FOR A BALANCE OF INTERESTS

Abstract. The article explores the fundamental conflict between the needs of modern artificial intelligence (AI) systems, particularly large language models (LLMs), for vast amounts of data ("data hunger") and personal data protection standards.

The relevance of the research is driven by the rapid adoption of generative AI (ChatGPT, Gemini, Copilot), which creates systemic risks to privacy, including algorithmic bias, deanonymization, and

the opacity of "black boxes". Technological innovations are significantly outpacing the development of adequate legal regulation.

The purpose of the article is to examine the legal, technical, and organizational measures being taken to find a balance between the interests of AI developers and the rights of data subjects.

The article analyzes: 1) the incompatibility of key principles of the General Data Protection Regulation (data minimization, purpose limitation) with the architecture of large language models (LLMs) and the complementary role of the EU AI Act; 2) the effectiveness of technical measures (PETs), such as synthetic data, trusted execution environments, and federated learning; 3) the organizational approaches and privacy policies of OpenAI, Google, and Microsoft.

The conclusions emphasize that companies have implemented a dichotomy in protection guarantees: robust contractual terms for corporate clients and an "opt-out" model for mass consumers, which shifts the burden of privacy protection onto the user. The implementation of technical measures remains partial due to their resource-intensive nature. It is concluded that existing measures are insufficient to ensure a real balance of interests for individual users.

Keywords: artificial intelligence; personal data; General Data Protection Regulation (GDPR); EU AI Act; large language models (LLMs); privacy policy; ChatGPT; Gemini; Copilot.