ПРАВО ЗАРУБІЖНИХ КРАЇН

DOI: https://doi.org/10.31617/3.2026(142)09
УДК 342.9:004.8

ГАЧКЕВИЧ Андрій
https://orcid.org/0000-0002-8494-1937

к. ю. н., доцент,
доцент кафедри міжнародного
та кримінального права
Інституту права, психології
та інноваційної освіти
Національного університету "Львівська
політехніка"

вул. Степана Бандери, 12, м. Львів,
79013, Україна
andrii.o.hachkevych@lpnu.ua

HACHKEVYCH Andrii
https://orcid.org/0000-0002-8494-1937

PhD (Law), Associate Professor,
Associate Professor at the Department
of International and Criminal Law
Institute of Law, Psychology, and
Innovation Education
Lviv Polytechnic National University

12, Stepan Bandera St., Lviv,
79013, Ukraine
andrii.o.hachkevych@lpnu.ua

"СІРІ ЗОНИ" ПРАВОВОГО РЕГУЛЮВАННЯ ШІ ТА ОБРОБКИ ПЕРСОНАЛЬНИХ ДАНИХ

У сучасний період штучний інтелект часто застосовують для виконання завдань, які так чи інакше пов'язані з обробкою персональних даних. Відповідно до чинного законодавства будь-які операції з персональними даними повинні здійснюватись відкрито, прозоро, а також пропорційно визначеним цілям. Водночас для систем штучного інтелекту характерний феномен "чорної скриньки", який полягає в тому, що спосіб їхнього функціонування є малозрозумілим та зумовлює ситуації, які можуть не мати чіткого правового вирішення, – вони відбуваються у "сірій зоні". Розглянуто, чому проблема захисту персональних даних набуває особливої гостроти внаслідок розвитку штучного інтелекту та суміжних технологій. В основу дослідження покладено гіпотезу, що відсутність прозорості стосовно неможливості пояснення суб'єктові даних, яку саме інформацію про нього та у який спосіб система штучного інтелекту буде використовувати, є єдиним фактором порушення захисту персональних даних. Для того щоб її перевірити, проаналізовано поширені способи застосування штучного інтелекту: профілювання, біометричну ідентифікацію, генерування контенту, а також процедуру навчання систем штучного інтелекту,

"GREY AREAS" OF LEGAL REGULATION OF AI AND PERSONAL DATA PROCESSING

In the modern period, artificial intelligence is often used to perform tasks that are in one way or another related to the processing of personal data. According to current legislation, any operations with personal data must be carried out openly, transparently, and in proportion to defined purposes. At the same time, AI systems are characterized by the "black box" phenomenon, which consists in the fact that the way they function is poorly understood and causes situations that may not have a clear legal solution – they occur in a "grey area". It is considered why the issue of personal data protection is becoming particularly acute due to the development of artificial intelligence and related technologies. The study is based on the hypothesis that the lack of transparency regarding the impossibility of explaining to the data subject what specific information about him and in what way the AI system will use it is the only factor violating the personal data protection. To test this, common applications of AI were analyzed: profiling, biometric identification, content generation, as well as the procedure of training AI systems, which is associated with web scraping as a method of collecting information from websites. Since the

що пов'язана з вебскрейпінгом як збором інформації з вебсайтів. Оскільки наведені способи у своїй більшості позначені неологізмами, за допомогою етимологічного методу описано їхню сутність. Вивчено поняття обробки персональних даних у законодавстві держав світу та показано, за якими критеріями щодо об'єкта обробки можна оцінювати способи обробки на предмет потенційної небезпечності. Результати дослідження приводять до висновку, що обробка персональних даних системами штучного інтелекту може становити неприйнятний ризик, порушувати вимогу точності персональних даних, створювати загрози для посиленого захисту "чутливих даних" та інформації про дітей, відбуватись без поінформованості суб'єктів даних, а також породжувати ситуацію, коли персональні дані зберігаються значно довше, ніж потрібно для визначеної цілі, та коли їх неможливо виправити (реалізувати право на виправлення). Фактори, які спричиняють порушення захисту персональних даних, крім відсутності прозорості, включають неможливість забезпечення повного контролю над оперативними процесами штучного інтелекту, активну взаємодію з іншими системами і мережами, а також недостатню захищеність від витоків даних, несанкціонованого доступу та інших загроз.

Ключові слова: штучний інтелект, захист персональних даних, обробка персональних даних, штучний інтелект та персональні дані, профілювання, вебскрейпінг.

mentioned methods are mostly designed by neologisms, their essence is described using the etymological method. The concept of personal data processing in the legislation of countries around the world has been studied, and it has been shown according to which criteria regarding the processing object one can assess processing methods for potential danger. The results of the study lead to the conclusion that the processing of personal data by AI systems may pose an unacceptable risk, violate the requirement for personal data accuracy, create threats to the enhanced protection of "sensitive data" and information about children, occur without the awareness of data subject, and also create a situation where personal data is stored significantly longer than necessary for the defined purpose and when it cannot be corrected (exercise the right to rectification). The factors that cause violations of personal data protection, aside from the lack of transparency, include the inability to ensure full control over the AI operational processes, active interaction with other systems and networks, as well as insufficient protection against data leaks, unauthorized access, and other threats.

Keywords: artificial intelligence; personal data protection; personal data processing; artificial intelligence and personal data; profiling; web scraping.

JEL Classification: K10, K19, K20, K29, K33.

Вступ

Розвиток штучного інтелекту та суміжних технологій має значний вплив на відносини щодо обробки персональних даних – системи штучного інтелекту використовують персональні дані для виконання різного роду завдань: від розпізнавання облич як способу біометричної ідентифікації до демонстрування рекомендацій користувачам соціальних мереж.

При цьому сфера обігу персональних даних підпорядкована строгому правовому регулюванню: будь-які операції з персональними даними повинні здійснюватися відповідно до положень спеціального законодавства, частиною якого є принципи обробки даних. Серед них – прозорість – "обробка персональних даних здійснюється відкрито і прозоро із застосуванням засобів та у спосіб, що відповідають визначеним цілям такої обробки" (Закон України "Про захист персональних даних" № 2297-VI, 2010). Принцип прозорості вимагає забезпечення зрозумілості тих процесів, які відбуваються в системах штучного інтелекту. Щоправда, відсутність прозорості, через яку суб'єкт даних не отримує всієї інформації щодо того, як система штучного інтелекту використовує інформацію про нього, може бути далеко не єдиним фактором порушення захисту персональних даних.

Відповідно, ця стаття присвячена проблемі захисту персональних даних з урахуванням тих серйозних викликів, які породжує штучний інтелект та суміжні технології на сучасному етапі.

У дослідженнях правових аспектів використання штучного інтелекту проблема захисту персональних даних займає важливе місце, що підтверджують праці таких українських вчених, як В. Базалицький, А. Бакуменко, Д. Бєлов, М. Бєлова, А. Гачкевич, В. Демчина, Н. Загребельна, В. Некрутенко, Є. Остіян, А. Саміло, М. Суржинський та ін.

Ключова ідея, яка співзвучна зі змістом пропонованої статті, полягає в тому, що і приватність, і конфіденційність як цінності потребують особливої охорони, оскільки перебувають у постійній небезпеці в епоху штучного інтелекту і суміжних технологій.

Водночас необхідність прийняття краще адаптованого до сучасних реалій законодавства за зразком Загального регламенту про захист даних ЄС, забезпечення його ефективного дотримання, насамперед під наглядом спеціально створеної інституції, пошук організаційних і технічних можливостей для того, щоб охорона приватності та конфіденційності, а також захист персональних даних були більш надійними, – це питання, які порушували такі українські автори: І. Бем, М. Блохін, І. Городиський, Н. Головацький, І. Дейкун, О. Легка, І. Машкіна, Я. Овчаренко, С. Рзаєва, П. Складанний та ін.

В іноземній науці взаємозв’язок штучного інтелекту та захисту персональних даних є популярним напрямом дослідження, представленим величезною кількістю публікацій. У центрі уваги часто перебуває складне співвідношення технологічного прогресу і питання персональних даних. Наприклад, *Humerick* (2018) вважає, що у ЄС суворі вимоги Загального регламенту про захист даних аж ніяк не сприяють навчанню систем штучного інтелекту та їхньому вдосконаленню. У комплексному дослідженні впливу Загального регламенту про захист даних на штучний інтелект *Sartor* та *Lagioia* (2020) дійшли висновку, що контролерам слід впроваджувати відповідальний підхід до обробки персональних даних. Необхідною є й співпраця всіх зацікавлених сторін – контролерів даних, урядових інституцій, наглядових органів та організацій громадянського суспільства – для того щоб сприяти законності обробки персональних даних. *Onik et al.* (2019), вважаючи персональні дані "новою нафтою" або "новим полем військових дій" у контексті Четвертої промислової революції, пояснюють, в чому полягають небезпеки сучасних технологій і штучного інтелекту для персональних даних, включно з несанкціонованим доступом та витоками даних.

Фокус уваги цієї статті спрямований на способи застосування штучного інтелекту, які з точки зору охорони приватності та конфіденційності викликають занепокоєння через ситуації, що можуть не мати чіткого правового вирішення, а отже, відбуватись у "сірій зоні".

В основу дослідження покладено гіпотезу щодо визначення відсутності прозорості операцій систем штучного інтелекту – суб’єкти даних не можуть отримати повних пояснень, яка інформація про них та

у який спосіб буде використана, що є єдиним фактором порушення захисту персональних даних. Для її перевірки проведено обраний метою статті розгляд окремих способів застосування штучного інтелекту, а також процедури навчання систем штучного інтелекту, що пов'язана з вебскрейпінгом як збором інформації з вебсайтів.

Зважаючи на те, що предмет цього дослідження – обробка персональних даних системами штучного інтелекту – пов'язаний із поняттям обробки, насамперед потрібно з'ясувати його значення відповідно до законодавства сучасних держав. Далі показано, за якими критеріями стосовно об'єкта способи обробки можуть бути оціненими на предмет потенційної небезпечності. За допомогою етимологічного методу описано окремі способи застосування штучного інтелекту: профілювання, біометричну ідентифікацію, генерування контенту, а також навчально-орієнтовану практику вебскрейпінгу. З використанням системного підходу узагальнено фактори неналежної охорони приватності та конфіденційності, а також порушення захисту персональних даних.

1. Поняття обробки персональних даних у законодавстві сучасних держав та критерії потенційної небезпечності залежно від об'єкта обробки

Понятійно-категоріальний апарат законів світових держав про захист персональних даних містить термін "обробка персональних даних". У *табл. 1* наведено його визначення з метою відображення різних типів національних правових систем.

Таблиця 1

Поняття обробки персональних даних у законодавстві держав світу

Держава або штат (назва закону)	Визначення обробки персональних даних
Алжир (Закон про захист фізичних осіб у сфері обробки персональних даних від 10.06.2018)	Будь-яка операція або сукупність операцій, що здійснюються як з використанням автоматизованих засобів та процесів, так і без нього, щодо персональних даних, зокрема: збір, запис, упорядкування, зберігання, адаптування або зміна, пошук, ознайомлення, використання, розкриття шляхом передачі, поширення або надання іншим способом, зіставлення або поєднання, а також блокування, шифрування, стирання або знищення
Бразилія (Закон про захист персональних даних від 14.08.2018)	Будь-яка операція, що виконується з персональними даними, як-от збір, виробництво, отримання, класифікація, використання, доступ, відтворення, передача, розповсюдження, обробка, подання, зберігання, видалення, оцінка інформації або контроль, модифікація, повідомлення, передача, поширення або пошук
Південна Африка (Закон про захист персональної інформації від 19.11.2013)	Будь-яка операція або діяльність, будь-який набір операцій, що здійснюються з використанням автоматизованих засобів або без них, щодо персональних даних, включно зі: (а) збором, отриманням, записом, організацією, зіставленням, зберіганням, оновленням або модифікацією, пошуком, зміною, консультацією або використанням; (б) поширенням шляхом передачі, розповсюдженням або наданням доступу в будь-якій іншій формі; (в) об'єднанням, пов'язуванням, а також обмеженням, погіршенням якості, стиранням або знищенням інформації

Закінчення табл. 1

Держава або штат (назва закону)	Визначення обробки персональних даних
Каліфорнія (Закон про приватність споживачів від 28.06.2018)	Будь-яка операція або сукупність операцій, що виконуються з персональними даними або наборами персональних даних, незалежно від того, чи виконуються вони автоматизованими засобами чи ні
Сінгапур (Закон про захист персональних даних від 20.11.2012)	Виконання будь-якої операції або сукупності операцій стосовно персональних даних, що включає: запис; збереження; організацію, адаптацію або зміну; пошук; комбінацію; передачу різними способами; стирання або знищення
Україна (Закон про захист персональних даних від 01.06.2010)	Будь-яка дія або сукупність дій, як-от збирання, реєстрація, накопичення, зберігання, адаптування, зміна, поновлення, використання і поширення (розповсюдження, реалізація, передача), знеособлення, знищення персональних даних, у тому числі з використанням інформаційних (автоматизованих) систем

Джерело: складено автором на основі опрацювання понятійно-категоріального апарату законодавства про персональні дані сучасних держав: Алжиру (*Loi relative à la protection des personnes physiques dans le traitement des données à caractère personnel*, 2018), Бразилії (*Brazilian Data Protection Law LGPD*, 2018), Південної Африки (*No. 4 of 2013: Protection of Personal Information Act*, 2013), Каліфорнії (*California Legislative Information,* 2018), Сінгапуру (*Singapore Statutes Online*, 2012), України (Закон України "Про захист персональних даних" № 2297-VI, 2010).

Варто додати, що в Загальному регламенті про захист даних "обробка персональних даних" означає будь-яку операцію або сукупність операцій з персональними даними або наборами персональних даних, незалежно від того, чи вони виконуються автоматизованими засобами чи ні. До таких операцій належать: збір, запис, організація, структурування, зберігання, адаптація або зміна, пошук, консультація, використання, розкриття шляхом передачі, поширення або забезпечення доступності іншим способом, узгодження або поєднання, обмеження, видалення або знищення (*GDPR*, 2013).

На основі вивчення поняття обробки в законодавстві сучасних держав, включно з державами – членами ЄС, закони яких узгоджені з положеннями Загального регламенту про захист даних, можна зауважити, що воно охоплює чотири види операцій із персональними даними:

- збір – коли система штучного інтелекту отримує інформацію про фізичну особу і персональні дані стають доступними для використання;
- використання – за допомогою персональних даних як ресурсу, здатного до змін, сучасні технології виконують ті чи інші завдання, отримують певні результати;
- зміна якості – з одного боку, персональні дані можуть бути виправленими, тим самим ставати більш точними, з іншого – зі

зниженням рівня якості персональні дані втрачають чутливість, наслідком десенсибілізації є й підвищення рівня захищеності;

- видалення – коли інформація перестає існувати в системі штучного інтелекту, а отже, не підлягає подальшому використанню.

Дотримання принципів обробки персональних даних та забезпечення прав суб'єкта даних як гарантій правомірності обробки повинно здійснюватись з того моменту, коли системи штучного інтелекту починають накопичувати інформацію про фізичну особу будь-яким методом (якщо тільки вона вже не стала доступною з інших підстав), та завершуватись з видаленням даних після досягнення визначеної цілі.

Водночас не до кінця зрозуміло, чи можливо видалити інформацію, яка стала доступною для системи штучного інтелекту. Риторичний характер цього питання, як і питання того, чи справді можливо виправити неточні персональні дані, свідчить про те, що технічні можливості на сучасному етапі не відповідають прийнятим стандартам.

Варто зазначити, що в положеннях більшості проаналізованих у межах дослідження законів поряд із персональними даними виокремлена категорія "чутливі персональні дані" (інколи – спеціальні), яка має посилений захист: операції з такими даними вимагають особливих гарантій для суб'єктів даних (ст. 11 Закону Бразилії звужує підстави для обробки порівняно зі звичайними персональними даними), а в деяких випадках обробка може бути взагалі практично забороненою (ст. 22 Загального регламенту про захист даних стосовно автоматизованого прийняття рішень).

Слід додати, що залежно від сфер застосування штучного інтелекту відносини з обробки персональних даних, крім загальних правил, можуть бути підпорядковані спеціальним (наприклад стосовно генетичних даних). Такий підхід проявляється в розробці рекомендацій галузевого характеру, внесенні змін до наявних законодавчих актів або навіть прийнятті нових, як-от Закону Каліфорнії про приватність генетичної інформації (*California Legislative Information*, 2021).

Зауважимо наявність у законодавстві сучасних держав ще однієї категорії персональних даних із посиленим захистом – про дітей (ст. 34–35 Закону Південної Африки).

Для оцінювання потенційної небезпечності обробки персональних даних системами штучного інтелекту доцільно змоделювати своєрідний спектр, від позиції у якому залежить необхідність вжиття заходів зі зменшення ризиків.

Критерії щодо об'єкта обробки, які показують приналежність відповідної операції до того чи іншого кінця спектра, є такими: наявність "чутливих персональних даних" або звичайних; інформація

стосується дітей або не пов'язана з ними; персональні дані представляють один із виокремлених у законодавстві видів чи є загальними (в деяких державах наразі не існує законодавчого розрізнення видів персональних даних). Якщо об'єктом операції є персональні дані, які за кожним із критеріїв є спеціальними, то такі операції ближче до того кінця спектра, який відображає особливу небезпечність.

При застосуванні систем штучного інтелекту три вищенаведені категорії персональних даних як об'єкт обробки здебільшого не виділяються із загалу – повний контроль над оперативними процесами користувачами систем наразі не може бути забезпеченим через специфіку сучасних технологій. Крім того, для фільтрування даних за переліченими критеріями перед основним використанням штучний інтелект поки не запрограмований, хоча він безумовно спроможний на це.

2. Способи застосування штучного інтелекту: профілювання, біометрична ідентифікація, генерування контенту

У переліку операцій, які набули поширення сьогодні та під час яких персональні дані підлягають обробці, варто розглянути профілювання, біометричну ідентифікацію, а також генерування контенту.

Профілювання походить від слова "профіль" – означає сукупність основних типових рис, які характеризують господарство, фах тощо (Дмитрієв, 2019*b*). Поняття профілювання є дуже розлогим за змістом та може охоплювати як широко застосовуваний персоналізований маркетинг, так і формування психологічних портретів у некомерційних цілях – правопорушника для боротьби зі злочинністю або оцінку ментального стану пацієнта, щоб надати правильне лікування.

Профілювання – це спосіб застосування штучного інтелекту, а також форма обробки персональних даних, коли з окремих характеристик суб'єкта даних система штучного інтелекту робить висновки про його приналежність до певної категорії. При цьому профілювання – це не тільки класифікування, а й метод передбачити поведінку суб'єкта даних – за положеннями Загального регламенту про захист даних – для аналізу або прогнозування аспектів, що стосуються ефективності в роботі, економічного становища, здоров'я, особистих уподобань, інтересів, надійності, поведінки, місцезнаходження або пересувань (*GDPR*, 2013).

Як уже зазначалося, профілювання є дуже поширеним у цифровому маркетингу – воно полягає у демонструванні особі реклами, яка швидше за все її зацікавить (завдяки попередньо отриманій інформації про цю особу, навіть у малому обсязі, можна передбачити її поведінку як споживача).

До профілювання мав відношення скандал щодо витоку даних користувачів *Facebook* у 2015–2017 рр. за участю компанії *Cambridge*

Analytica. Він показав, наскільки високою є чутливість персональних даних у процесах аналітики даних з використанням штучного інтелекту. Інформація з мільйонів профілів у *Facebook* зазнала витоку даних та незаконної обробки в цілях політичного профілювання (*Kleinman*, 2018, *March* 21).

Слід зауважити, що з прийняттям Закону ЄС про штучний інтелект профілювання як метод оцінки ризиків у контексті боротьби зі злочинністю виведений із "сірої зони" до заборонених практик: використання систем штучного інтелекту для оцінки ризиків, пов'язаних з фізичними особами, з метою оцінки або прогнозування можливості вчинення кримінального правопорушення, виходячи лише з профілювання фізичної особи або оцінки її особистісних рис і характеристик, є протиправним (*European Union*, 2025).

Зі все більш активним впровадженням біометрична ідентифікація породжує дискусії з приводу того, наскільки вона відповідає "букві та духу" законодавства про захист персональних даних і дотриманню прав людини загалом. Безумовно, як процес, під час якого штучний інтелект може визначити, ким є та чи інша людина, біометрична ідентифікація – біометричний означає такий, що належить до деталізованої інформації про тіло людини, наприклад про особливості кольору її очей, яка може бути використана для підтвердження особи цієї людини (*Cambridge Dictionary*, 2025*a*) – є необхідною для кількох важливих цілей, при цьому має відбуватися з урахуванням поваги до цінностей приватності та конфіденційності.

Автоматичне розпізнавання особи за її унікальними біологічними характеристиками (відбитками пальців, звучанням голосу, скануванням райдужної оболонки ока, скануванням обличчя, геометрією мочки вуха) сприяє підвищенню ефективності правоохоронної діяльності, а також автоматизації й оптимізації контролю над міграційними та іншими процесами у сфері державного управління. Біометрична ідентифікація порушує питання балансу: на одній шальці терезів – інтереси національної безпеки та боротьби зі злочинністю, на іншій – дотримання прав людини при використанні систем штучного інтелекту.

Водночас із розвитком штучного інтелекту обробка біометричних даних застосовуватиметься все частіше і приватними компаніями (наприклад для контролю за робочим часом працівника). Вже сьогодні такого роду ідентифікація інтегрована в системи розумного будинку та для авторизації цифрових пристроїв, а технології розпізнавання облич, розроблені *Cleaview AI* після збору величезних обсягів фотографій у мережі "Інтернет", викликали не тільки дискусії в наукових колах, на які вплинула закупівля продуктів від *Clearview AI* для діяльності правоохоронних органів, а й пред'явлення правових претензій у різних державах (*Borak*, 2025, *June* 11).

Відповідно до ст. 5 Закону ЄС про штучний інтелект через неприйнятний ризик біометрична ідентифікація на основі штучного інтелекту в багатьох випадках є забороненою:

- створення або розширення бази даних розпізнавання облич шляхом нецільового збору зображень облич з інтернету або записів із камер відеоспостереження;
- розпізнавання емоцій фізичної особи на робочому місці та в навчальних закладах, за винятком медичних або безпекових підстав;
- визначення приналежності до окремої раси, ймовірної підтримки певних політичних поглядів, релігійних чи філософських переконань, ймовірного членства в профспілці, а також особливостей статевого життя чи сексуальної орієнтації (за винятками систематизації законно отриманих наборів біометричних даних і категоризації для цілей правоохоронної діяльності);
- використання систем дистанційної біометричної ідентифікації в режимі реального часу в загальнодоступних місцях для цілей правоохоронної діяльності, водночас у виключному порядку ця процедура таки може здійснюватись відповідно до визначених у Законі ЄС про штучний інтелект правил (*European Union*, 2025).

"Унікальна ідентифікація людини" підпадає під особливий захист з огляду на дію відповідних положень законодавства у сфері обігу персональних даних, зокрема стосовно автоматизованого прийняття рішень у контексті ст. 9 Загального регламенту про захист даних (*Bulgakova*, 2022).

Під генеруванням контенту маємо на увазі створення штучним інтелектом умовно нової інформації – текстів, включно з комп'ютерним кодом, зображеннями, аудіо та відео. Чому умовно нової? Адже вихідні дані є результатом складних процесів, що відбуваються всередині "чорної скриньки", а не в людській свідомості. Творчість генеративних систем штучного інтелекту важко назвати повноцінною хоча б через відсутність творчого внеску людини, тому, вочевидь, більш доречно вживати поняття "синтетична творчість" (*Hachkevych*, 2025*a*).

У спрощеному розумінні генерування контенту – слово "контент" означає інформаційне наповнення (Дмитрієв, 2019*a*) – відбувається завдяки тому, що системи штучного інтелекту, опрацювавши величезні обсяги даних, навчились за виявленими шаблонами та взявши за основу користувацькі налаштування створювати: фотографії, які показують реальних людей, хоча ніколи не були зроблені насправді; звукові файли, коли відомі виконавці співають пісні інших відомих виконавців, хоч насправді ніколи їх не виконували, або романи, написані в стилі видатних письменників (імітуючи їхню манеру).

Зрештою, найбільш поширеною формою генерування контенту є відповіді розмовних чат-ботів (*ChatGPT*, *Microsoft Copilot*, *Gemini* тощо) на питання користувачів.

Серед викликів застосування розмовних чат-ботів – введення або завантаження до систем штучного інтелекту інформації, яка містить персональні дані, а також виведення ними результатів з такою ж особливістю (*Hachkevych*, 2025*b*).

Відносно нещодавно стало зрозумілим, що обробка персональних даних розмовними чат-ботами, як і іншими системами штучного інтелекту, наприклад для резюмування новинних повідомлень, може призвести до порушення принципу точності: сервіс *Apple Intelligence* повідомив, що начебто корпорація *BBC* оприлюднила новину про самогубство Л. Манджоне, заарештованого після вбивства генерального директора *UnitedHealth*, що не відповідало дійсності (*Fraser*, 2024, *December* 13); М. Бернклау, німецький журналіст, виявив, що чат-бот *Microsoft Copilot* надавав недостовірну інформацію про нього – як про втікача з психіатричної лікарні, шахрая, наркоторговця та жорстокого злочинця (ймовірно, штучний інтелект взяв до уваги ті справи, про які писав журналіст) (*Kelsey-Sugg & Carrick*, 2024, *November* 3); Б. Худ, австралійський посадовець, був названий *ChatGPT* співучасником вчинення кримінального правопорушення, тоді як насправді він був одним із викривачів (*Mayers et al.*, 2023, *April* 6).

3. Небезпеки процедури навчання систем штучного інтелекту для персональних даних, а також інші ймовірні ризики, зумовлені обробкою персональних даних

Навчання систем штучного інтелекту має на меті зробити їх спроможними виконувати ті завдання, для яких вони призначені. Процедура навчання полягає в тому, що алгоритми аналізують величезні обсяги інформації, в тому числі й персональні дані, для того щоб виявити наявні закономірності – завдяки їм штучний інтелект буде здатним видавати очікувані результати: приймати рішення, прогнозувати, рекомендувати тощо. Надалі система штучного інтелекту зможе опрацьовувати будь-які вхідні дані, адже вона знайшла потрібний алгоритм.

Вебскрейпінг, також дата-скрейпінг, на комп’ютерному жаргоні означає діяльність, що передбачає отримання інформації з вебсайту або екрану комп’ютера та внесення її до електронної таблиці, тобто електронного документа, в якому інформація міститься у рядках і стовпцях та підлягає використанню для розрахунків (*Cambridge Dictionary*, 2025*b*). Вебскрейпінг полягає в тому, що спеціальне програмне забезпечення за короткий час витягує різного роду інформацію з вебсайтів, включно з персональними даними. Такі дані потрібні системам штучного інтелекту з низки причин, вони є цінним навчальним ресурсом та формують свого роду експертність у технологій.

В оцінках правомірності вебскрейпінгу єдності поки не існує – з одного боку, інформація, на яку націлений вебскрейпінг, є загальнодоступною, тому вона може бути ресурсом для навчання систем штучного інтелекту. Навіть якщо раніше сталися витік персональних даних та незаконне оприлюднення інформації про фізичну особу в мережі "Інтернет", то причина цього зовсім не в зборі даних системою штучного інтелекту. Крім того, навчання відрізняється від застосування, а отже, шкода від того, що над персональними даними буде проведено автоматизований аналіз для виявлення загальних закономірностей, зовсім не співрозмірна розкриттю конфіденційної інформації для широкого кола осіб або її необмеженому в часі зберіганню.

З іншого боку, навіть за умови загальнодоступності персональні дані не повинні втрачати правового захисту. Певні паралелі можна навести з авторським правом: оприлюднення в інтернеті твору, який підлягає правовій охороні, не дозволяє вільно використовувати його. Для такого використання має бути отримана згода правовласника або, як мінімум, наявні правові підстави щодо вільного використання твору без згоди автора. Відповідно, для того щоб практики вебскрейпінгу не відбувались у "сірій зоні", слід керуватись положеннями законодавства для регулювання відносин у сфері обігу персональних даних. У цьому аспекті найбільш прогресивним є Загальний регламент про захист даних, що визначає принципи обробки, права суб'єктів даних, а також формальні підстави, за яких обробка є законною.

Слід зазначити, що питання правомірності обробки персональних даних при навчанні систем штучного інтелекту було порушеним Європейською радою з захисту даних (*European Data Protection Board*, 2024), а також у судовій практиці, зокрема в рішенні Вищого земельного суду Кельна (*Oberlandesgericht Köln*, 2025). Відповідно, при розробці та вдосконаленні систем штучного інтелекту персональні дані можуть бути використаними за наявності законного інтересу як формальної підстави для обробки даних, що підтверджується за допомогою триетапної перевірки:

(1) визначення, в чому полягає законний інтерес для контролера даних або третьої сторони;

(2) проведення аналізу необхідності обробки для забезпечення визначеного законного інтересу (так званий тест на необхідність);

(3) оцінювання того, чи законний інтерес не переважає інтереси або основні права та свободи суб'єктів даних ("тест на балансування").

Вищий земельний суд Кельна, своєю чергою, постановив, що використання персональних даних користувачів *Facebook* та *Instagram* для навчання систем штучного інтелекту корпорацією *Meta* не є протиправним: навіть за відсутності згоди користувачів на обробку їхніх персональних даних існує формальна підстава для оброрбки у вигляді законного інтересу.

Обробка персональних даних системами штучного інтелекту пов’язана і з іншими ризиками, які проявляються, зокрема, в таких сферах застосування штучного інтелекту, як електронна комерція, охорона здоров’я, а також інформаційні технології. Наявність наведених далі ризиків у *табл. 2* додатково підкреслює той факт, що обробка персональних даних системами штучного інтелекту належить до "сірої зони" та потребує вдосконалення правового вирішення.

Таблиця 2

Ризики обробки персональних даних
системами штучного інтелекту в різних сферах

Сфера використання штучного інтелекту	Ймовірні ризики, зумовлені обробкою персональних даних
Електронна комерція	Витік даних. Ризики та небезпеки, пов’язані з третіми сторонами (наприклад компанією, яка надає маркетингові послуги). Згода на обробку даних не була отримана належним чином (механізм *cookies* не налаштований). Зберігання даних про користувачів протягом необмеженого строку, тобто порушення принципу пропорційності визначеній цілі
Медичні послуги	Несанкціонований доступ, зокрема внаслідок "слабких" паролів. Втрата даних та атаки програм-вимагачів. Порушення функціонування порталів для пацієнтів (ризики підвищуються при неправильних налаштуваннях). Людська помилка, яка може бути результатом недостатнього рівня цифрових навичок або необізнаності персоналу
Інформаційні технології	Транскордонна передача даних. Профілювання, за якого аналіз даних здійснюється непрозоро. Ризики та небезпеки на рівні самої компанії, включно з необмеженим доступом працівників до файлів користувачів. Масштабованість процесів, яка випереджає вжиття заходів безпеки

Джерело: адаптовано автором для цілей цього дослідження на основі рекомендацій стосовно управління ризиками (*General Data Protection Regulation*, 2025, *April* 9).

Висновки

Характерний для штучного інтелекту феномен "чорної скриньки", зумовлений тим, що спосіб його функціонування є малозрозумілим, а отже, прозорість – як передумова пояснюваності будь-яких операцій з персональними даними – не може бути забезпеченою, певною мірою визначає зв’язок обробки персональних даних системами штучного інтелекту з "сірою зоною".

Як наслідок того, що невідомою є інформація про дані, які використовує штучний інтелект, про те, як він ці дані опрацьовує та у який спосіб приймає рішення, необхідні для виконання різного роду

завдань, до сучасних технологій виникає недовіра, яку посилюють інші виявлені та досліджувані у цій статті фактори.

По-перше, використання систем штучного інтелекту може породжувати ризики – від прийнятних, якими інколи можна знехтувати, до неприйнятних, через високу ймовірність яких експлуатація систем взагалі може бути заборонена (наприклад для соціального скорингу або профілювання в цілях притягнення до кримінальної відповідальності).

Водночас і способи обробки персональних даних можуть перебувати в різних кінцях спектра потенційної небезпечності: інформація про дітей, "чутливі персональні дані" або генетичні дані як об'єкти обробки несуть високий ризик, їхня допустимість залежить від запровадження ефективних заходів для мінімізації потенційної шкоди.

По-друге, після того як інформація у будь-який спосіб стала доступною для системи штучного інтелекту, точність як вимога часто не може бути виконаною. Операції, які відбуваються всередині відповідно до свого роду логіки штучного інтелекту, слабо підконтрольні користувачам, особливо тим, які не мають спеціальних знань та навичок. Відповідно, навіть якщо дані в первісному наборі будуть виправленими, це не змінить кінцевого результату, що добре помітно в контексті генерування контенту.

По-третє, обробка даних може відбуватись без поінформованості суб'єктів даних, особливо для навчання систем штучного інтелекту. Без попередження користувачів та отримання їхньої згоди штучний інтелект може зберігати та використовувати введену ними інформацію, в тому числі персональні дані із завантажених документів.

Зауважимо, що дедалі частіше, згідно з прийнятими стандартами захисту персональних даних, загальнодоступні сервіси штучного інтелекту, включно з розмовними чат-ботами, містять опцію надання згоди на те, щоб інформація, яка передається системі, надалі сприяла навчанню штучного інтелекту. Водночас те, що інформація про фізичну особу була оприлюднена в мережі "Інтернет", не дозволяє внаслідок вебскрейпінгу вільно здійснювати її збір та використання.

Однією з найбільших перешкод для того, щоб охорона приватності та конфіденційності, а також захист персональних даних при використанні систем штучного інтелекту були належними, є те, що персональні дані зберігаються значно довше, ніж потрібно для визначеної цілі. Крім того, наявні сприятливі умови для обміну даними, витоків даних та несанкціонованого доступу – програма на базі штучного інтелекту взаємодіє з іншими програмами. Хоча сама по собі система штучного інтелекту є окремим технічним рішенням, розробленим для виконання певного завдання, вона інтегрована в технологічне середовище і може бути під'єднана до інтернету та інших мереж.

У процесі дослідження спростовано гіпотезу про те, що єдиним фактором порушення захисту персональних даних є відсутність прозорості стосовно неможливості пояснення суб'єктові даних, яку саме інформацію про нього та у який спосіб система штучного інтелекту буде використовувати. Отримані результати показують, що при обробці

персональних даних системами штучного інтелекту такі дані можуть втрачати свою якість, водночас виправити це поки неможливо, а також можуть зберігатись значно довше, ніж потрібно для визначеної цілі. Повною мірою не забезпечується посилений захист "чутливих даних", інформації про дітей, генетичних та інших виокремлених у законодавстві видів персональних даних, як і повний контроль людини над процесами, що відбуваються всередині систем штучного інтелекту.

У зв'язку з цим напрями подальших досліджень визначає необхідність вдосконалювати технічні й організаційні аспекти штучного інтелекту, поряд із правовими, щоб персональні дані, які він обробляє, могли бути убезпеченими, а також зміненими в разі необхідності або навіть видаленими цілком після того, як заявлена мета була реалізована.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ / REFERENCES

Borak, M. (2025, June 11). *Clearview AI faces more legal uncertainty in UK and US*. Biometric Update. https://www.biometricupdate.com/202506/clearview-ai-faces-more-legal-uncertainty-in-uk-and-us

Brazilian Data Protection Law LGPD. (2018). ANPD. https://www.gov.br/anpd/pt-br/centrais-de-conteudo/outros-documentos-e-publicacoes-institucionais/lgpd-en-lei-no-13-709-capa.pdf

Bulgakova, D. (2022). Unique human identification under the GDPR Article 9(1) (2). *Philosophy of Law and General Theory of Law*, *1*, 130–159. https://doi.org/10.21564/2707-7039.1.275645

California Legislative Information. (2018). *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

California Legislative Information. (2021). *California Genetic Information Privacy Act of 2021.* https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202120220SB41

Cambridge Dictionary. (2025a). *Biometric.* https://dictionary.cambridge.org/dictionary/english/biometric

Cambridge Dictionary. (2025b). *Web-scraping.* https://dictionary.cambridge.org/dictionary/english/web-scraping

European Data Protection Board. (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models.* https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

European Union. (2025). *The EU Artificial Intelligence Act*. The Artificial Intelligence Act. https://artificialintelligenceact.eu/

Fraser, G. (2024, December 13). *BBC complains to Apple over misleading shooting headline.* BBC. https://www.bbc.com/news/articles/cd0elzk24dno

GDPR. (2013). *Art. 4 – Definitions*. In *General Data Protection Regulation (GDPR)*. https://gdpr-info.eu/art-4-gdpr/

General Data Protection Regulation. (2025, April 9). *Best 3 GDPR risk assessment examples.* https://gdprinfo.eu/best-3-gdpr-risk-assessment-examples

Hachkevych, A. (2025a). "Synthetic Creativity" of Generative Artificial Intelligence Poses Challenges for Legal Protection of Copyright and Related Rights. *Bulletin of Lviv Polytechnic National University.* Series: Legal Sciences, *12*(3), 43–50. https://doi.org/10.23939/law2025.47.043

Hachkevych, A. (2025b). The impact of generative artificial intelligence on the legal systems of contemporary states. *Law and Innovative Society, 1*(24), 37–46. https://doi.org/10.37772/2309-9275-2025-1(24)-3

Humerick, M. (2018). Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence. *Santa Clara High Technology Law Journal, 34*(4), 393–418.

Kelsey-Sugg, A., & Carrick, D. (2024, November 3). *AI hallucinations caused artificial intelligence to falsely describe these people as criminals*. ABC News. https://www.abc.net.au/news/2024-11-04/ai-artificial-intelligence-hallucinations-defamation-chatgpt/104518612

Kleinman, Z. (2018, March 21). *Cambridge Analytica: The story so far.* BBC News. https://www.bbc.com/news/technology-43465968	
Loi relative à la protection des personnes physiques dans le traitement des données à caractère personnel. (2018). *Journal Officiel de la République Algérienne Démocratique et Populaire, Conventions et Accords Internationaux.* https://www.joradp.dz/FTP/JO-FRANCAIS/2018/F2018034.pdf	Law relating to the protection of natural persons in the processing of personal data. (2018). *Official Journal of the People's Democratic Republic of Algeria, Conventions and International Agreements*. https://www.joradp.dz/FTP/JO-FRANCAIS/2018/F2018034.pdf
Mayers, L., Martin, S., & Rybicki, D. (2023, April 6). *Victorian mayor may sue OpenAI after ChatGPT "accuses" him in bribery case.* ABC News. https://www.abc.net.au/news/2023-04-06/hepburn-mayor-flags-legal-action-over-false-chatgpt-claims/102195610	
No. 4 of 2013: Protection of Personal Information Act, 2013. (2013). Republic of South Africa Government Gazette. https://www.gov.za/sites/default/files/gcis_document/201409/3706726-11act4of2013protectionofpersonalinforcorrect.pdf	
Oberlandesgericht Köln, 15 UKl 2/25. (2025). Justiz.nrw.de. https://nrwe.justiz.nrw.de/olgs/koeln/j2025/15_UKl_2_25_Urteil_20250523.html	*Higher Regional Court of Cologne, 15 UKl 2/25*. (2025). Justiz.nrw.de. https://nrwe.justiz.nrw.de/olgs/koeln/j2025/15_UKl_2_25_Urteil_20250523.html
Onik, M. M. H., Kim, C.-S., & Yang, J. (2019). Personal Data Privacy Challenges of the Fourth Industrial Revolution. In *2019 21st International Conference on Advanced Communication Technology (ICACT)* (pp. 635–638). IEEE.	
Sartor, G., & Lagioia, F. (2020). *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. European Parliament. https://doi.org/10.2861/293	
Singapore Statutes Online. (2012). *Personal Data Protection Act 2012 – Singapore Statutes Online*. https://sso.agc.gov.sg/Act/PDPA2012	
Дмитрієв, О. (2019a). *Контент – тлумачення, орфографія, новий правопис онлайн.* https://slovnyk.ua/index.php?swrd=контент	Dmytriiev, O. (2019a). *Content – interpretation, spelling, new online orthography.* https://slovnyk.ua/index.php?swrd=контент
Дмитрієв, О. (2019b). *Профіль – тлумачення, орфографія, новий правопис онлайн.* https://slovnyk.ua/index.php?swrd=профіль	Dmytriiev, O. (2019b). *Profile – interpretation, spelling, new online orthography.* https://slovnyk.ua/index.php?swrd=профіль
Закон України "Про захист персональних даних" № 2297-VI (2010). Офіційний вебпортал парламенту України. https://zakon.rada.gov.ua/laws/show/2297-17#Text	The Law of Ukraine "On the Protection of Personal Data" No. 2297-VI. (2010). Official Website of the Parliament of Ukraine. https://zakon.rada.gov.ua/laws/show/2297-17#Text

Конфлікт інтересів. Автор заявляє, що не має фінансових чи нефінансових конфліктів інтересів щодо цієї публікації; не має відносин із державними органами, комерційними або некомерційними організаціями, які могли б бути зацікавлені у поданні цієї точки зору.

Автор не отримував прямого фінансування для цього дослідження.

Гачкевич, А. (2026). "Сірі зони" правового регулювання ШІ та обробки персональних даних. *Ius Modernum*, *1*(142), 121–135. https://doi.org/10.31617/3.2026(142)09

Надійшла до редакції 23.09.2025.
Прийнято до друку 29.10.2025.
Публікація онлайн 12.03.2026.