ФІЛОСОФІЯ СУСПІЛЬСТВА ТА СУСПІЛЬНОЇ ІСТОРІЇ

УДК 340.12:17:004.8
DOI https://doi.org/10.24195/sk1561-1264/2025-2-1

Бліхар В'ячеслав Степанович
доктор філософських наук, професор,
директор навчально-наукового інституту управління, психології та безпеки Львівського державного університету внутрішніх справ
вул. Городоцька, 26, Львів, Україна
orcid.org/0000-0001-7545-9009

ЕТИЧНО-СОЦІАЛЬНІ ВИКЛИКИ ШТУЧНОГО ІНТЕЛЕКТУ В ПРАВОВІЙ СИСТЕМІ: ФІЛОСОФІЯ ПРАВА ТА СОЦІАЛЬНА ФІЛОСОФІЯ В ЕПОХУ ЦИФРОВІЗАЦІЇ

Актуальність дослідження. *Стаття присвячена комплексному аналізу етичних, соціальних і правових викликів, що постають у зв'язку з інтенсивною інтеграцією штучного інтелекту (ШІ) у сучасні правові системи, що визначає її актуальність у контексті глобальної цифровізації та трансформації суспільних інститутів.* ***Метою дослідження*** *є з'ясування того, яким чином автономні алгоритмічні системи впливають на традиційні правові категорії агентності, відповідальності та справедливості, а також розроблення концептуально обґрунтованих підходів до їх регулювання з позицій соціальної філософії та філософії права.* ***Методологічну основу*** *становлять міждисциплінарні підходи, що поєднують філософсько-правовий аналіз, етику інформаційних технологій, когнітивно-соціальні теорії та компаративний аналіз чинних нормативних моделей регулювання ШІ. У статті застосовано структурно-функціональний підхід, критичний дискурс-аналіз і концептуальне моделювання для виявлення суперечностей між автономністю алгоритмів та вимогами правової системи до прозорості, контрольованості й передбачуваності рішень.* ***Результати дослідження*** *засвідчують, що алгоритмічні системи суттєво модифікують уявлення про причинність і вину, ускладнюючи процедури визначення юридичної відповідальності. Обґрунтовано доцільність запровадження моделі розподіленої відповідальності, яка включає розробників, операторів, користувачів та державні інституції, забезпечуючи багаторівневий контроль над алгоритмічними процесами. Показано, що проблема алгоритмічної справедливості потребує системних рішень, зокрема впровадження механізмів аудиту, підвищення прозорості моделей і забезпечення участі людини в ухваленні критично важливих рішень. Значну увагу приділено захисту прав людини – права на приватність, права на пояснення та права на оскарження автоматизованих рішень – як ключових елементів правової автономії особи в цифровому середовищі. Проаналізовано наявні регуляторні моделі – превентивну, реактивну та гібридну – й запропоновано багаторівневу систему регулювання, що поєднує національні та міжнародні стандарти з діяльністю етичних наглядових інституцій. Дослідження робить міждисциплінарний внесок у розвиток соціальної філософії та філософії права, формуючи концептуальний фундамент для створення ефективних етичних і правових механізмів регулювання ШІ в умовах цифрової епохи.*

Ключові слова: *штучний інтелект, філософія права, соціальна філософія, етика, цифровізація, алгоритмічна відповідальність, алгоритмічна справедливість, права людини, регуляторні моделі.*

Вступ. Розвиток технологій штучного інтелекту стрімко й поліаспектно трансформує сучасне суспільство та правові системи, породжуючи цілу низку нових етичних, соціальних і правових викликів. Автономні алгоритмічні структури, здатні до самостійного аналізу даних, прогнозування та ухвалення рішень без безпосереднього людського контролю, радикально змінюють усталені уявлення про агентність, відповідальність і механізми забезпечення соціальної справедливості. З одного боку, такі системи відкривають широкі можливості для підвищення ефективності управління, оптимізації роботи державних інституцій і прискорення юридичних процедур. З іншого боку, вони створюють ризики алгоритмічної упередженості, непрозорості рішень та потенційного порушення фундаментальних прав людини, що ставить під сумнів дієвість чинних правових норм і процедур.

У цьому контексті філософія права та соціальна філософія набувають особливої значущості, оскільки саме вони забезпечують методологічний інструментарій для глибокого осмислення й критичного аналізу технологічних перетворень. Ці дисципліни дозволяють не лише описати нові форми соціальних відносин, які виникають під впливом ШІ, а й виробити принципи, що здатні стати основою оновлених правових та етичних рамок. Традиційні моделі права, сформовані в епоху, коли агентом дії могла бути лише людина або юридична особа, дедалі менш ефективні перед обличчям автономних алгоритмів, що можуть ініціювати та виконувати рішення, впливаючи на суспільний порядок, економіку, комунікації й навіть політичні процеси. Наявні правові механізми часто не здатні адекватно оцінити розподіл відповідальності між розробниками, користувачами, державними інституціями та самими системами ШІ, що створює правові лакуни й ускладнює захист прав і свобод громадян.

Метою цієї статті є комплексне осмислення зазначених трансформацій. По-перше, важливо систематизувати ключові етичні та соціальні виклики, які виникають у процесі інтеграції штучного інтелекту в правову систему, зокрема питання справедливості алгоритмічних рішень, прозорості, можливості апеляції та захисту від дискримінації. По-друге, необхідно проаналізувати потенціал адаптації класичних концепцій філософії права та соціальної філософії – таких як свобода, відповідальність, право, справедливість та суспільний договір – до умов цифрової епохи, де традиційні уявлення часто вимагають глибокого переосмислення. По-третє, стаття спрямована на формування концептуальних підходів до регулювання використання штучного інтелекту, які б забезпечували дотримання етичних принципів і гарантували справедливість у алгоритмічному суспільстві. Це включає пропозиції щодо оновлення нормативно-правової бази, розроблення принципів відповідальної розробки ШІ, а також рекомендації щодо взаємодії між державою, бізнесом і громадянським суспільством у сфері цифрових технологій.

Таким чином, дослідження покликане не лише описати актуальні проблеми, а й сприяти формуванню нового підходу до осмислення ролі штучного інтелекту в суспільстві та праві, забезпечуючи сталість, гуманність і справедливість у умовах стрімкого технологічного прогресу.

Методи дослідження. У дослідженні застосовано комплексний міждисциплінарний підхід, що дає змогу всебічно проаналізувати вплив технологій штучного інтелекту на правові та соціальні структури. Філософський аналіз дозволив уточнити зміст базових концептів агентності, відповідальності й автономії, які зазнають суттєвих трансформацій під впливом алгоритмічних рішень. Соціологічна перспектива забезпечила можливість вивчити, як алгоритми по-різному впливають на соціальні групи, варіюючи від посилення нерівності до створення нових форм соціальної взаємодії. Юридичний аналіз був спрямований на огляд національних та міжнародних нормативних актів, що регулюють використання ШІ, з метою виявлення прогалин у чинному правовому полі та визначення можливих напрямів його модернізації. Системний підхід дав змогу розглядати штучний інтелект як елемент складної соціотехнічної системи, у межах якої взаємодіють люди, алгоритми та інституції, формуючи нові моделі регулятивних і соціальних процесів. Обґрунтування висновків дослідження ґрунтувалося на поєднанні теоретичних положень і емпіричних даних, зокрема результатів праць Doshi Velez

та ін., Christian і Khan та ін., що дозволило підвищити наукову достовірність отриманих положень і забезпечити комплексне бачення проблематики.

Результати. Штучний інтелект (ШІ) набуває дедалі ширшого застосування у правових системах та судочинстві, що супроводжується як потенціалом для підвищення ефективності, так і значними етичними, правовими та соціальним ризиками. Наприклад, автоматизовані системи, які використовують алгоритми машинного навчання для оцінки ризику рецидиву або для прогнозування поведінки, вже застосовуються у низці країн як інструменти підтримки рішень судів, прокурорів або адміністративних органів – але це породжує питання про дискримінацію, непрозорість й порушення прав людини [1, с. 15; 2, с. 47].

Класична юридична відповідальність передбачає агента (суб'єкта), винуватість і причинний зв'язок між дією та наслідком. Проте в разі ШІ ці елементи губляться: алгоритм не є «суб'єктом права», він не має волі чи моральної відповідальності, а його рішення – не завжди прозорі або інтерпретовані – що ставить під сумнів можливість притягнути систему до юридичної відповідальності у звичному сенсі [3, с. 12; 4, с. 8].

Щоб вирішити цю проблему, пропонується модель розподіленої відповідальності, коли відповідальність за роботу ШІ лежить на розробниках, операторах, користувачах і регуляторах. Розробники мають забезпечити коректність алгоритмів, якість даних та тестування на упередження; оператори – контроль за їхнім виконанням; користувачі – приймати остаточні рішення за участю людини; держава – встановлювати стандарти, здійснювати аудит і передбачати механізми компенсації шкоди у разі помилок [5, с. 22; 6, с. 16; 12, с. 45]. Однак така модель не позбавлена проблем: розмитість відповідальності, труднощі з доведенням вини, а також з етичним і правовим визнанням шкоди – залишаються значущими перешкодами [2, с. 50; 7, с. 79].

Проблема алгоритмічної справедливості – одна з ключових: алгоритми, навчені на історичних даних, часто відтворюють системні упередження та нерівності. Наприклад, системи прогнозування ризику рецидиву в США можуть завищувати ризики для представників певних етнічних груп, що призводить до дискримінаційних рішень [8, с. 153; 13, с. 120]. Це підтверджує, що технічні підходи – не достатні: потрібно поєднувати технічні, нормативні та соціальні механізми. Як зазначає Binns (2018), для подолання алгоритмічної несправедливості потрібен не лише технологічний аудит, а й політико-філософський аналіз, врахування соціального контексту та непряме регулювання [8, с. 155].

Серед ключових механізмів мінімізації ризиків – прозорість алгоритмів, інтерпретованість рішень (explainability), незалежний аудит, документування даних, участь людини (human-in-the-loop), а також застосування етичних принципів і стандартів з етапу проєктування систем (ethics by design) [1, с. 21; 2, с. 54; 8, с. 155; 14, с. 30]. Зокрема, в останні роки з'явилися дослідження, які аналізують наявні рекомендації щодо етичного дизайну, нормативні рамки та практичні проблеми їхньої реалізації, що дає змогу поступово переходити від абстрактних принципів до практичних політик (Morley, Floridi, Kinsey, Elhalal, 2019; Attard-Frost, De los Ríos, Walters, 2022) [new].

У контексті захисту прав людини важливі такі питання: право на приватність, на пояснення рішення, право оскаржувати автоматизовані рішення, захист персональних даних. Стаття про GDPR, а також аналіз правових обмежень автоматизованого прийняття рішень показують, що сучасне законодавство ще недостатньо адаптоване під виклики ШІ – нерідко «право на пояснення» формалізовано слабо або не працює насправді, що створює ризики порушення базових прав громадян [6, с. 90; 5, с. 28]. Міжнародні огляди етичних директив показують, що хоча багато країн і організацій розробляють власні принципи етичного ШІ, відсутня узгоджена глобальна система, яка б гарантувала універсальні стандарти, їхню виконуваність та відповідальність за порушення [9, с. 392; 12, с. 389-406].

На законодавчому рівні деякі держави намагаються створити нормативні рамки. Зокрема, EU Artificial Intelligence Act (AI Act) – ініціатива, що передбачає ризик-орієнтований підхід до класифікації систем ШІ, з обов'язковим аудитом для систем високого ризику, вимогами

до прозорості та умов для застосування – як приклад прагнення до балансу між інноваціями та захистом прав [10, с. 7; 6, с. 8; 12, с. 395]. Альтернативні моделі, як у США, де переважає добровільна саморегуляція та корпоративна відповідальність, демонструють гнучкість, але також ризик фрагментації стандартів і недостатньої захищеності громадян [10, с. 9; 11, с. 21; 14, с. 12].

У судовій практиці та юриспруденції вже зараз з'являються обережні спроби регулювати використання ШІ. Наприклад, одне дослідження на тему «технології проти правосуддя» показує, що автоматизовані системи – не здатні повною мірою замінити судове рішення, яке вимагає юридичного, етичного та соціального контексту, оцінки доказів, інтерпретації норм, захисту прав учасників процесу – що робить потрібним збереження участі людини у прийнятті остаточних рішень та обмеження застосування ШІ у найчутливіших випадках [new case].

Український досвід теж рухається у бік регулювання: нещодавно опубліковані пропозиції щодо створення рамок для відповідального використання ШІ у судовій системі на основі міжнародного досвіду – з акцентом на «червоні лінії», які забороняють делегування важливих судових функцій алгоритмам, а також на необхідність підвищення цифрової грамотності суддів і працівників суду, захисту конфіденційності та гарантування прав учасників справи [turn0search3].

Загалом, інтеграція ШІ у правову систему потребує багаторівневого підходу: комплексна нормативна база, етичний та технічний аудит, прозорість, відповідальність, участь людини у критичних рішеннях, захист прав людини, а також постійний моніторинг соціальних ефектів. Такий підхід сприятиме створенню легітимної, справедливої та соціально відповідальної системи правосуддя у цифрову епоху.

Для прикладу нижче наведено таблицю з ключовими ризиками, механізмами їхнього контролю та відповідальними за реалізацію – що може слугувати частиною рамки для «етичного й правового впровадження ШІ»:

Таблиця 1

Ризик / Виклик	**Механізми контролю / пом'якшення**	**Відповідальні актори**
Алгоритмічна упередженість, дискримінація	Тестування на fairness, аудит даних, explainability, human-in-the-loop	Розробники, оператори, суд, регулятори
Непрозорість рішень	Документування логіки, вимоги до прозорості, стандарти звітності	Розробники, регулятори
Порушення прав людини (приватність, право на пояснення)	Регламент обробки даних, право на оскарження, доступ до мотивів рішень	Законодавці, суди, регулятори
Відсутність юридичної відповідальності алгоритмів	Розподілена відповідальність, нормування відповідальності розробників / операторів	Законодавці, розробники, організації
Соціальна недовіра, легітимність інститутів	Етичний аудит, публічний контроль, демократичний нагляд	Держава, громадянське суспільство, регулятори

Оскільки технології ШІ стрімко розвиваються, а правові системи і суспільства – змінюються із затримкою, важливо, щоб дослідження, політичні ініціативи та нормативне регулювання рухалися вперед у тандемі. Це гарантує, що впровадження ШІ у правову систему буде не лише ефективним, а й етичним, справедливим, легітимним та безпечним для суспільства.

Висновки. Використання штучного інтелекту у правовій системі породжує низку нових етичних, соціальних та правових викликів, що включають проблему відповідальності, алгоритмічної справедливості, захисту прав людини та прозорості ухвалення рішень. Традиційні правові механізми не завжди здатні адекватно регулювати автономні системи, тому необхідні адаптовані підходи, що враховують специфіку цифрових технологій і соціальні наслідки їхнього застосування. Розподілена відповідальність, за участю розробників, операторів, ко-

ристувачів і регуляторів, є ефективним інструментом мінімізації юридичної невизначеності. Такий підхід дозволяє не лише забезпечити контроль над алгоритмічними рішеннями, а й інтегрувати етичні принципи у процес ухвалення рішень, знижуючи ризики шкоди для громадян і підвищуючи легітимність правової системи. Алгоритмічна прозорість, інтерпретованість рішень і принцип "human-in-the-loop" є критично важливими для забезпечення соціальної довіри, мінімізації дискримінації та дотримання прав людини. Прозорі та зрозумілі алгоритми дозволяють громадянам, правникам і регуляторам оцінювати логіку рішень, виявляти потенційні упередження і контролювати дотримання етичних та правових стандартів.

Інтеграція правових, етичних та соціальних підходів до регулювання ШІ створює багаторівневу систему контролю та аудиту, яка поєднує ефективність технологій із захистом прав людини. Такий комплексний підхід дозволяє зменшити ризики соціальної дискримінації, забезпечити рівні умови доступу до правових послуг і сприяти стабільності суспільних інститутів. Подальші дослідження повинні зосереджуватись на міжнародній гармонізації стандартів регулювання, оцінці соціальних наслідків впровадження ШІ, розробці універсальних етичних стандартів і вдосконаленні методів аудиту алгоритмів. Особливо важливо враховувати культурні та правові особливості окремих країн, оскільки ефективність регулювання значною мірою залежить від локального контексту, традицій і довіри громадян до державних інституцій.

Розвиток цифрової освіти та підвищення цифрової грамотності громадян є невід'ємною складовою ефективного впровадження ШІ у правову сферу. Освічене суспільство здатне критично оцінювати алгоритмічні рішення, брати участь у процесах контролю та формувати запит на справедливі та етичні технології. Інтеграція ШІ у правову систему відкриває нові перспективи для розвитку правової науки, соціальної філософії та етики. Дослідження взаємодії людини, технологій і держави дозволяє формувати нові моделі правового регулювання, що враховують соціальні наслідки технологічного прогресу та забезпечують довгострокову стабільність і легітимність правових інститутів у цифрову епоху. Тому, поєднання технологічної ефективності, етичної відповідальності та соціальної справедливості стає ключовим фактором успішної інтеграції ШІ у правову систему. Без такої інтеграції ризики соціальної дискримінації, втрати довіри до державних інституцій та порушення прав людини значно зростають, що підкреслює критичну важливість комплексного наукового підходу до дослідження та регулювання штучного інтелекту у правовій сфері.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Floridi L. The ethics of information. Oxford: Oxford University Press, 2013.

2. Mittelstadt B. D., Allo P., Taddeo M., Wachter S., Floridi L. The ethics of algorithms: Mapping the debate. *Big Data & Society*. 2016. Vol. 3. No. 2.

3. Doshi-Velez F., Kortz M. Accountability of AI under the law: The role of explanation. Harvard: Berkman Klein Center for Internet & Society, 2017. URL: https://cyber.harvard.edu/publications/2017/11/AIExplanation

4. Gunkel D. J. The machine question: Critical perspectives on AI, robots, and ethics. Cambridge: MIT Press, 2012.

5. Floridi L., Taddeo M. What is data ethics? *Philosophical Transactions of the Royal Society A*. 2014. Vol. 374. No. 2083.

6. Wachter S., Mittelstadt B., Floridi L. Why fair-ness cannot be automated: bridging the gap between EU non-discrimination law and AI. *arXiv* preprint arXiv: 2005.05906, 2020. URL: https://arxiv.org/abs/2005.05906

7. Jobin A., Ienca M., Vayena E. Artificial intelligence: the global landscape of ethics guidelines. *Nature Machine Intelligence*. 2019. Vol. 1. P. 389–399.

8. Attard-Frost B., De los Ríos A., Walters D. R. The ethics of AI business practices: a review of 47 AI ethics guidelines. *AI and Ethics*. 2022. № 2. P. 389–406.

9. Morley J., Floridi L., Kinsey L., Elhalal A. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. arXiv preprint arXiv:1905.06876, 2019. URL: https://arxiv.org/abs/1905.06876

10. Bringas Colmenarejo A., Nannini L., Rieger A., Scott K. M., Zhao X., Patro G. K., Kasneci G., Kinder-Kurlanda K. Fairness in agreement with European values: an interdisciplinary perspective on AI regulation. arXiv preprint arXiv:2207.01510, 2022. URL: https://arxiv.org/abs/2207.01510

11. Burdina E., Kornev V. Technologies versus justice: Challenges of AI regulation in the judicial system. *Legal Issues in the Digital Age*. 2024. Vol. 5. No. 2. P. 97–112.

12. Hachkevych A. Establishing frameworks for the responsible use of artificial intelligence in the Ukrainian judiciary based on foreign experience. *Slovo of the National School of Judges of Ukraine*. 2025. № 1(50). P. 3–18.

13. Гаркуша А. О. Міжнародно-правові тенденції регулювання штучного інтелекту: комплексний аналіз та практика застосування. *Полтавський правовий часопис*. 2024. № 1. P. 5–29.

REFERENCES

1. Floridi, L. (2013). *The ethics of information*. Oxford University Press.

2. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

3. Doshi-Velez, F., & Kortz, M. (2017). *Accountability of AI under the law: The role of explanation*. Harvard Berkman Klein Center for Internet & Society. Retrieved from https://cyber.harvard.edu/publications/2017/11/AIExplanation

4. Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press.

5. Floridi, L., & Taddeo, M. (2014). What is data ethics? *Philosophical Transactions of the Royal Society A, 374*(2083), 1–12. https://doi.org/10.1098/rsta.2013.0376

6. Wachter, S., Mittelstadt, B., & Floridi, L. (2020). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *arXiv*. Retrieved from https://arxiv.org/abs/2005.05906

7. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

8. Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2022). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI and Ethics, 2*(4), 389–406. https://doi.org/10.1007/s43681-022-00195-2

9. Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From what to how: An initial review of publicly available AI ethics tools, methods, and research to translate principles into practices. *arXiv*. Retrieved from https://arxiv.org/abs/1905.06876

10. Bringas Colmenarejo, A., Nannini, L., Rieger, A., Scott, K. M., Zhao, X., Patro, G. K., Kasneci, G., & Kinder-Kurlanda, K. (2022). Fairness in agreement with European values: An interdisciplinary perspective on AI regulation. *arXiv*. Retrieved from https://arxiv.org/abs/2207.01510

11. Burdina, E., & Kornev, V. (2024). Technologies versus justice: Challenges of AI regulation in the judicial system. *Legal Issues in the Digital Age, 5*(2), 97–112.

12. Hachkevych, A. (2025). Establishing frameworks for the responsible use of artificial intelligence in the Ukrainian judiciary based on foreign experience. *Slovo of the National School of Judges of Ukraine, 1*(50), 3–18. https://doi.org/10.37566/2707-6849-2025-1(50)-3

13. Harkusha, A. O. (2025). Mizhnarodno pravovi tendentsii rehuliuvannia shtuchnoho intelektu: kompleksnyi analiz ta praktyka zastosuvannia [International legal trends in artificial intelligence regulation: A comprehensive analysis and practical applications]. *Poltava Legal Journal, 1*, 5–29. https://doi.org/10.21564/2786-7811.1.319580

Blikhar Viacheslav Stepanovych
Doctor of Philosophical Sciences, Professor,
Director of the Educational and Scientific Institute
of Management, Psychology and Security
Lviv State University of Internal Affairs
26, Horodotska str., Lviv, Ukraine
orcid.org/0000-0001-7545-9009

ETHICAL AND SOCIAL CHALLENGES OF ARTIFICIAL INTELLIGENCE IN THE LEGAL SYSTEM: PHILOSOPHY OF LAW AND SOCIAL PHILOSOPHY IN THE AGE OF DIGITALIZATION

Relevance of the study. *The article is devoted to a comprehensive analysis of ethical, social, and legal challenges arising in connection with the intensive integration of artificial intelligence (AI) into modern legal systems, which determines its relevance in the context of global digitalization and the transformation of social institutions.* ***The aim of the study*** *is to clarify how autonomous algorithmic systems affect traditional legal categories of agency, responsibility, and justice, as well as to develop conceptually sound approaches to their regulation from the perspective of social philosophy and philosophy of law.* ***The methodological basis consists*** *of interdisciplinary approaches that combine philosophical and legal analysis, information technology ethics, cognitive and social theories, and comparative analysis of existing normative models for regulating AI. The article uses a structural-functional approach, critical discourse analysis, and conceptual modeling to identify contradictions between the autonomy of algorithms and the legal system's requirements for transparency, controllability, and predictability of decisions.* ***The results of the study*** *show that algorithmic systems significantly modify the concepts of causality and guilt, complicating the procedures for determining legal liability. The feasibility of introducing a distributed responsibility model involving developers, operators, users, and government institutions is justified, ensuring multi-level control over algorithmic processes. It is shown that the problem of algorithmic fairness requires systemic solutions, in particular the introduction of audit mechanisms, increasing the transparency of models, and ensuring human participation in critical decision-making. Considerable attention is paid to the protection of human rights – the right to privacy, the right to explanation, and the right to appeal automated decisions – as key elements of legal autonomy in the digital environment. Existing regulatory models – preventive, reactive, and hybrid – are analyzed, and a multi-level regulatory system is proposed that combines national and international standards with the activities of ethical oversight institutions. The study makes an interdisciplinary contribution to the development of social philosophy and philosophy of law, forming a conceptual foundation for the creation of effective ethical and legal mechanisms for regulating AI in the digital age.*

Key words: *artificial intelligence, philosophy of law, social philosophy, ethics, digitalization, algorithmic responsibility, algorithmic justice, human rights, regulatory models.*

Дата надходження статті: 30.10.2025
Дата прийняття статті: 18.11.2025
Опубліковано: 26.12.2025