

УДК 340.12

DOI: <https://doi.org/10.24144/2307-3322.2026.94.1.10>

## АРХІТЕКТУРА ЦИФРОВОГО СУМЛІННЯ ЯК МЕХАНІЗМ ІМПЛЕМЕНТАЦІЇ АНТРОПОЦЕНТРИЗМУ ПРАВА В СИСТЕМИ ШТУЧНОГО ІНТЕЛЕКТУ

**Кравчук С. М.,**

*старший викладач кафедри теорії права та конституціоналізму  
Навчально-науковий інститут права, психології та інноваційної освіти  
Національний університет «Львівська політехніка»*

ORCID: 0000-0002-3485-1343

**Коваль А. В.,**

*здобувачка першого (бакалаврського) освітньо-кваліфікаційного рівня вищої  
освіти, студентка IV курсу, спеціальність 081 Право  
Навчально-науковий інститут права, психології та інноваційної освіти  
Національний університет «Львівська політехніка»*

ORCID: 0009-0002-3723-5105

**Кравчук С. М., Коваль А. В. Архітектура цифрового сумління як механізм імплементації антропоцентризму права в системи штучного інтелекту.**

Нині актуальним є доктринальне розв'язання проблеми збереження людської ідентичності та суб'єктивності в умовах тотальної алгоритмізації публічно-правових відносин. Стаття висвітлює проблему структурної дегуманізації права, що виникає внаслідок неконтрольованого делегування владних повноважень автоматизованим системам. Досліджується трансформація класичної антропоцентричної парадигми під тиском технологій штучного інтелекту, що призводить до формування феномену «алгоритмічного раціоналізму». Визначається необхідність переходу від ретроспективних моделей правового регулювання до проактивних стратегій забезпечення прав людини.

У статті обґрунтовується необхідність впровадження концепту «цифрового сумління» як інтегрованої в програмний код системи алгоритмічних запобіжників. В ході дослідження були використані такі методи як системно-структурний аналіз, формально-юридичний метод, метод компаративістики та прогностичного моделювання. Кризь призму аксіологічного підходу доведено, що антропоцентризм має стати базовим архітектурним імперативом при проектуванні цифрових екосистем. Особлива увага приділяється концепції «Law-by-Design», яка передбачає трансформацію конституційних принципів у технічні вимоги до нейронних мереж.

У статті аналізується сучасний стан законодавчого регулювання ШІ в Україні та Європейському Союзі, зокрема норми Акта про штучний інтелект та GDPR. Автори статті вважають, що існуючі етичні декларації є недостатніми без підкріплення жорсткими юридичними санкціями та протоколами алгоритмічного аудиту. Розглядається проблема «алгоритмічної упередженості» та прихованої дискримінації, що виникає внаслідок навчання моделей на стереотипізованих історичних даних. Виявлено колізії у правовому статусі «алгоритмічних агентів» у сфері публічних закупівель та соціального забезпечення.

У роботі зроблено висновок, що цифрова трансформація держави потребує створення незалежних органів нагляду за алгоритмічною підзвітністю. Особлива увага приділяється поняттям «цифрового соціального внеску» та «права людини на безперервну технологічну адаптацію» як відповіді на виклики технологічного безробіття. Встановлено доцільність імплементації принципу «Human-in-the-loop» у військовій та цивільній системах безпеки для гарантування персональної відповідальності за ухвалені рішення. Авторка Кравчук С. М. пропонує власні пропозиції щодо вдосконалення процесуального законодавства через запровадження презумпції «алгоритмічної вини» оператора системи.

**Ключові слова:** права людини, штучний інтелект, антропоцентризм, цифрове сумління, людська гідність, алгоритмічна упередженість, право на приватність, Human-in-the-loop, Legal-Code Hybridization.

**Kravchuk S. M., Koval A. V. The architecture of digital conscience as a mechanism for implementing anthropocentrism in artificial intelligence systems.**

A pressing issue today is the doctrinal resolution of the problem of preserving human identity and subjectivity in the context of the total algorithmicization of public-law relations. This article examines the problem of the structural dehumanization of law arising from the uncontrolled delegation of authority to automated systems. It examines the transformation of the classical anthropocentric paradigm under the pressure of artificial intelligence technologies, which leads to the emergence of the phenomenon of “algorithmic rationalism.” The article identifies the need to transition from retrospective models of legal regulation to proactive strategies for ensuring human rights.

The article justifies the need to introduce the concept of “digital conscience” as a system of algorithmic safeguards integrated into the program code. The study employed methods such as systemic-structural analysis, the formal-legal method, the comparative method, and predictive modeling. Through the lens of an axiological approach, it is demonstrated that anthropocentrism must become a fundamental architectural imperative in the design of digital ecosystems. Particular attention is paid to the concept of “Law-by-Design,” which entails the transformation of constitutional principles into technical requirements for neural networks.

The article analyzes the current state of AI regulation in Ukraine and the European Union, specifically the provisions of the AI Act and the GDPR. The authors of the article believe that existing ethical declarations are insufficient without being backed by strict legal sanctions.

The paper concludes that the digital transformation of the state requires the establishment of independent oversight bodies for algorithmic accountability. Particular attention is paid to the concepts of “digital social contribution” and “the human right to continuous technological adaptation” as responses to the challenges of technological unemployment. The paper establishes the feasibility of implementing the “human-in-the-loop” principle in military and civilian security systems to ensure personal accountability for decisions made. The author, S. M. Kravchuk, offers her own proposals for improving procedural legislation through the introduction of a presumption of “algorithmic fault” on the part of the system operator.

**Key words:** human rights, artificial intelligence, anthropocentrism, digital conscience, human dignity, algorithmic bias, right to privacy, human-in-the-loop, Legal-Code Hybridization.

**Постановка проблеми.** Стрімкий розвиток генеративних моделей та впровадження систем предиктивного аналізу в архітектуру державного управління створюють нову реальність, у якій класичні правові інститути опиняються на межі функціональної анахронічності. Проблема полягає у виникненні онтологічного розриву між статичною нормою права, орієнтованою на людську волю, та динамічним алгоритмом, що керується статистичними ймовірностями. За таких умов виникає реальна загроза перетворення людини з найвищої соціальної цінності на об’єкт цифрової класифікації, позбавлений дієвих механізмів захисту від автоматизованого свавілля.

Сучасна юриспруденція стикається з необхідністю не просто обмежити використання ШІ, а докорінно змінити саму філософію правотворення. Поширення концепцій трансгуманізму, що нівелюють духовну сутність особи, вимагає від держави формування активного імунітету у вигляді право-технічних інструментів захисту антропоцентризму. Це зумовлює потребу в дослідженні механізмів, які б дозволили інтегрувати правові імперативи безпосередньо в технологічний процес, роблячи порушення прав людини технічно неможливим.

З огляду на правовий режим воєнного стану в Україні, питання алгоритмічної прозорості набувають екзистенційного значення. Використання ШІ в оборонній сфері та при розподілі гуманітарних ресурсів загострює проблему відповідальності та справедливості. Відсутність чіткої архітектури «цифрового сумління» може призвести до системних помилок, що матимуть невідворотні наслідки для життя та безпеки мільйонів громадян, що робить дане дослідження вкрай нагальним для сучасної правової доктрини.

**Мета дослідження.** Метою дослідження є доктринальне обґрунтування концепції «цифрового сумління» як техніко-юридичного механізму імплементації антропоцентризму права в архітектуру

систем штучного інтелекту, а також формування практичних рекомендацій щодо подолання алгоритмічної упередженості в умовах трансформації правової системи України.

**Стан опрацювання проблематики.** Питанням впровадження систем штучного інтелекту в правове поле та реалізації принципу антропоцентризму у своїх наукових працях присвячували увагу відомі українські та зарубіжні вчені. Теоретичні засади антропоцентризму як ціннісного орієнтиру права ґрунтовно досліджувала Т. О. Подковенко. Проблематику дегуманізації цифровізації та ризику трансгуманізму висвітлювала у своїх розвідках Л. А. Пашко. Концептуальну платформу цифрового гуманізму як соціокультурної парадигми сучасності ідентифікувала І. Коломоєць.

Трансформацію політико-правових доктрин під впливом ШІ та європейські стандарти врядування аналізувала Т. О. Пікуля. Питання архітектури транзитивної державності та формування «алгоритмічної бюрократії» стали предметом наукового пошуку М. В. Сікала. Механізми правового регулювання цифрових процесів для стійкості економіки розкривав В. І. Міщенко. Глибоке переосмислення природи суб'єктивних прав особистості в умовах панування алгоритмів Big Data представлено у доробку С. М. Кравчук та С. Василенко. Дослідники не лише обґрунтували трансформацію класичних правових інститутів, а й запропонували інноваційні моделі соціального захисту, адаптовані до викликів цифрової епохи [8, с. 146-155].

Виклики для системи гарантій Європейської конвенції з прав людини та практику ЄСПЛ щодо біометричного нагляду вивчав І. М. Берназюк. Адміністративно-правові аспекти функціонування ШІ в системі ProZorro та правовий статус «алгоритмічних агентів» обґрунтував М. Сарнацький. Питання захисту свободи слова від дезінформації генеративного ШІ розглядала О. О. Карапетян, а аксіологічні аспекти використання військових алгоритмів – І. П. Марків.

**Вклад основного матеріалу.** Інтенсивний розвиток інформаційно-комунікаційних систем та повсюдне впровадження алгоритмів автоматизованого прийняття рішень актуалізують фундаментальну потребу доктринального переосмислення місця людини у технологізованому світі. Розбудова демократичного суспільства та правової держави нерозривно пов'язана із забезпеченням прав і свобод людини, що в сучасних умовах вимагає нових філософсько-правових підходів до розуміння природи цифрових трансформацій. Як слушно зазначає Т. О. Подковенко, право набуває рис основного ціннісного орієнтиру, фундаментальної суспільної засади, що виконує людинотворчу функцію і створює умови для реалізації індивідом його можливостей [1, с. 22].

Завдяки такому аксіологічному підходу розкриваються нові грані соціальної цінності права, яке не може існувати відокремлено від антропологічного виміру. Цей принцип знаходить своє найвище нормативне закріплення у статті 3 Конституції України, згідно з якою людина, її життя і здоров'я, честь і гідність, недоторканність і безпека визнаються найвищою соціальною цінністю, а утвердження і забезпечення прав і свобод людини є головним обов'язком держави [2].

В умовах стрімкої алгоритмізації суспільних відносин зазначений конституційний імператив постає перед безпрецедентними викликами, вимагаючи трансформації традиційних правових механізмів у нові форми захисту людської ідентичності.

Наше дослідження виходить із того, що цифровізація держави та суспільства неминує супроводжуватися ризиками структурної дегуманізації, які виникають унаслідок неконтрольованого делегування владних та управлінських повноважень автоматизованим системам. Дослідниця Л. А. Пашко аргументовано попереджає про небезпеку сучасних ідей трансгуманізму та постгуманізму, які базуються на відмові від врахування соціальних, культурних та духовних факторів існування людини, позиціонуючи технології як домінуючий елемент розвитку [3, с. 123].

Прихильники цих течій схильні розглядати людину лише як джерело даних для машинного навчання, нівелюючи її духовну та моральну сутність. На протигагу таким деструктивним тенденціям, єдиним ефективним запобіжником має стати концепція цифрового гуманізму. За визначенням І. Коломоєць, цифровий гуманізм є концептуальною платформою, що поєднує гуманістичні цінності та сучасні цифрові технології, імперативно визначаючи пріоритет людини в умовах технологічної трансформації суспільства [4, с. 60]. Відповідно, соціально-філософські засади цифрового гуманізму ґрунтуються на принципах людиноцентризму, автономії волі, етичної відповідальності та солідарності. Цифрові технології повинні підпорядковуватися цілям людського розвитку, а не перетворювати людину на об'єкт алгоритмічного управління чи елемент масиву великих даних.

Аналізуючи трансформацію політико-правових доктрин держави і права, неможливо ігнорувати той факт, що класичні теорії суспільного договору та державного суверенітету формувалися у виключно антропоцентричній парадигмі.

Т. О. Пікуля у своєму дослідженні ґрунтовно доводить, що класичні доктрини виходили з уявлення про державу як про інституціоналізовану форму публічної влади, засновану на волі людини, де лише людина визнавалася носієм політичної волі, розуму та відповідальності [5, с. 183].

Впровадження систем штучного інтелекту (далі – ШІ) у сферу публічного управління, правосуддя та соціальної політики докорінно порушує цю монополію. Виникає концептуально новий феномен «алгоритмічного раціоналізму», який базується не на загальних правових нормах чи моральних імперативах, а на статистичних кореляціях та математичних ймовірностях, видобутих із великих масивів даних. Це породжує гостре напруження між ідеєю правової визначеності, яка є основою верховенства права, та гнучкістю машинного навчання, що часто функціонує за принципом так званої «чорної скриньки».

Для глибшого розуміння структурних змін у механізмах публічного адміністрування доцільно звернутися до концепції «транзитивної моделі цифрової державності».

Як зазначає М. В. Сікало, цифрова трансформація спричиняє формування нової «алгоритмічної бюрократії», в межах якої значна частина рутинних, а подекуди й стратегічних рішень приймається автоматизованими системами на основі задалегідь встановлених правил та методів машинного навчання [6, с. 158]. Цей процес не є простою зміною формату документообігу; він являє собою перерозподіл владних повноважень між людськими та штучними агентами.

Ключовим аспектом імплементації цифрового сумління є фундаментальний перехід від традиційної парадигми правового регулювання до проактивних моделей інженерії. Дотепер юридична відповідальність та правове реагування здійснювалися переважно за принципом *ex-post* – тобто у формі покарання, компенсації або скасування незаконного акта вже після того, як правопорушення відбулося і шкоду було фактично завдано. В умовах швидкості, невидимості та масштабованості дій штучного інтелекту такий ретроспективний підхід виявляється катастрофічно неефективним, адже алгоритм здатний дискримінувати тисячі громадян за частки секунди.

Складність імплементації етико-правових норм у цифрове середовище детально аналізує В. І. Міщенко, який вказує на розробку міжнародними організаціями низки стандартів, зокрема Принципів Організації економічного співробітництва та розвитку та Рекомендацій Організації Об'єднаних Націй з питань освіти, науки і культури щодо етики штучного інтелекту [7, с. 285].

Проте, попри свою концептуальну значущість у формуванні ціннісних орієнтирів, ці документи здебільшого залишаються інструментами «м'якого права» (*soft law*), які покладаються на добровільне саморегулювання корпорацій і не містять жорстких інституційних механізмів державного примусу. Синтезуючи наведені підходи, ми доходимо висновку про безальтернативність переходу до принципу *ex-ante* регулювання, який на практиці втілюється у концепції «*Law-by-Design*» (право через проєктування).

Авторка даного наукового дослідження, Кравчук С. М., вважає, що антропоцентризм як архітектурний імператив права означає, що класичні конституційні принципи – повага до гідності, рівність перед законом, презумпція невинуватості, свобода волі – повинні перестати бути виключно текстами юридичних документів і трансформуватися у прямі «технічні завдання» для інженерів, архітекторів баз даних та розробників нейронних мереж. Право має диктувати системну архітектуру штучного інтелекту ще на самому початковому етапі написання базового коду. Якщо алгоритмічна модель демонструє ознаки прийняття дискримінаційних рішень або порушення приватності на етапі внутрішнього тестування, її подальше розгортання в публічному чи комерційному секторі має імперативно блокуватися на апаратному та програмному рівнях.

В наукових працях справедливо наголошується, що великомасштабні системи збору інформації здатні агрегувати величезні масиви персональних даних у реальному часі, створюючи вкрай деталізовані профілі поведінкових патернів, інтересів, пересувань та емоційних станів громадян без їхньої свідомої згоди [8, с. 149].

Використання технологій розпізнавання облич та предиктивного біометричного спостереження відкриває шлях до тотального цифрового нагляду, що є прямою загрозою основам демократичного суспільства та свободі волевиявлення. Фундаментальним правовим запобіжником у цьому контексті

служує Загальний регламент про захист даних (GDPR) [9], який встановлює жорсткі межі щодо збору, зберігання та автоматизованої обробки персональної інформації, імперативно вимагаючи дотримання принципів мінімізації даних, конфіденційності та цільового обмеження вже на етапі проєктування інформаційної системи (Privacy-by-Design).

Особливої уваги та детального вивчення заслуговує новаторський підхід європейського законодавця, втілений у Регламенті Європейського Парламенту та Ради щодо встановлення гармонізованих правил щодо штучного інтелекту (Акт про штучний інтелект) [10].

Аналізуючи цей епохальний документ, науковці констатують, що ризик-орієнтований підхід, запропонований Європейським Союзом, передбачає сувору диференціацію правового режиму виключно залежно від потенційного та реального впливу системи штучного інтелекту на фундаментальні права людини. Цей європейський акт прямо та безкомпромісно забороняє використання алгоритмічних систем з неприйнятним рівнем ризику, зокрема державні системи масового соціального скорингу або такі алгоритми, що використовують підсвідомі маніпулятивні техніки для викривлення поведінки людини. Для систем високого ризику запроваджуються жорсткі превентивні вимоги щодо прозорості логіки ухвалення рішень, гарантованої якості навчальних даних та обов'язкового постмаркетингового моніторингу. З доктринальної точки зору, ця прогресивна європейська концепція врядування формує нову модель гібридної державності, де класичні правові інститути демократії неподільно поєднуються з механізмами глибокого технологічного та алгоритмічного контролю.

На наше переконання, формування архітектури цифрового сумління вимагає рішучого подолання сліпого слідування оптимізаційним алгоритмам на шкоду індивідуальному людському підходу. Впровадження автоматизованих систем у сферу надання соціальних, медичних та адміністративних послуг, безперечно, експоненціально підвищує швидкість та формальну ефективність їх надання. Водночас, повне позбавлення алгоритму здатності до правової емпатії та соціальної чутливості призводить до жорстокого ігнорування складних життєвих обставин громадян, які фізично та юридично не піддаються простій бінарній комп'ютерній класифікації «так/ні». Відповідно, антропоцентризм права як вищий архітектурний імператив вимагає створення такого збалансованого цифрового середовища, де економічна чи часова ефективність алгоритмічної системи ніколи, за жодних обставин, не переважатиме над справедливістю ухваленого рішення щодо конкретної фізичної особи. Тільки за таких умов інтеграція штучного інтелекту в державне управління буде визнана легітимною та соціально прийнятною.

Логічним продовженням теоретичного осмислення антропоцентризму є пошук та формування дієвих, нормативно закріплених механізмів його впровадження у практичну площину діяльності держави. Сучасна доктрина публічного адміністрування та правознавства стикається із серйозним та безпрецедентним викликом: етичні декларації, меморандуми та маніфести, які домінують у міжнародному дискурсі щодо штучного інтелекту, щодня демонструють свою неспроможність зупинити системні порушення прав людини транснаціональними корпораціями та урядами.

Це свідчить про глибоку кризу ефективності норм «м'якого права». Відсутність жорстких юридичних санкцій, обов'язкових протоколів алгоритмічного аудиту та інститутів профільного державного нагляду призводить до того, що розробники використовують етичні кодекси радше як інструменти корпоративного піару та ухилення від реальної юридичної відповідальності, аніж як дійсні запобіжники проти «алгоритмічного свавілля». Відтак, на думку авторки даної статті Кравчук С. М., нагальною є потреба розробки та законодавчого закріплення концепції право-технічної гібридизації (Legal-Code Hybridization) – складного процесу прямої трансформації імперативних норм матеріального права у машиночитаний алгоритмічний формат, що стає невід'ємною та незмінною частиною архітектури «цифрового сумління» будь-якої публічної інформаційної системи.

Фундаментальною перепорою на шляху такої імплементації є проблема доктринального визначення правового статусу автоматизованих систем у публічних правовідносинах. М. Сарнацький у своєму дослідженні аргументовано обґрунтовує необхідність переосмислення статусу штучного інтелекту як функціонально релевантного суб'єкта адміністративної взаємодії, особливо у таких сенситивних та фінансово містких сферах, як публічні закупівлі [11, с. 130].

Чинне законодавство України, зокрема базовий Закон України «Про публічні закупівлі» [12], визначає єдиним джерелом управлінської волі виключно фізичну особу (уповноважену особу замовника), повністю ігноруючи об'єктивний факт масштабного та щоденного залучення систем машинного

аналізу до оцінки, автоматичного ранжування та відхилення тендерних пропозицій на електронних майданчиках.

Поглиблюючи дослідження цифрових закупівельних процесів, варто звернути увагу на аналіз електронної системи ProZorro, яка є фундаментом транспарентності державних фінансів в Україні.

Як зазначає М. Сарнацький, впровадження ризик-індикаторів та інтеграція технологій ШІ в електронні системи публічних закупівель відкриває широкі перспективи для автоматизованого моніторингу та прогнозування недоброчесної поведінки учасників [13, с. 81]. Проте, використання алгоритмізованих механізмів без належного та детального правового регулювання створює низку серйозних юридичних загроз: від упереджених або дискримінаційних рішень щодо відсіювання постачальників до фактичного обмеження доступу учасників до справедливого правосуддя у разі помилкової автоматичної оцінки їхньої тендерної документації. Усунення цієї масивної прогалини вимагає термінового законодавчого визнання нової правової категорії – «алгоритмічного агента» – із чітким, вичерпним визначенням меж його технічної компетенції, обов'язком розкриття логіки функціонування та встановленням механізмів солідарної юридичної відповідальності людських суб'єктів, які впровадили його в експлуатацію.

Окремого комплексного розгляду в цьому контексті потребує системна проблема «алгоритмічної упередженості» (algorithmic bias) та явища «цифрової дискримінації», яке набуває прихованих форм. Конституційне право на рівність та абсолютну заборону дискримінації є однією з фундаментальних засад вільного суспільства. Дискримінація в епоху глибокого машинного навчання стає «невидимою», системною та високоавтоматизованою [14, с. 111].

Традиційне адміністративне та цивільне право спрямовані на боротьбу з усвідомленою, суб'єктивною дискримінаційною дією конкретної фізичної чи юридичної особи (наприклад, упередженого керівника). Натомість, штучний інтелект продукує упередженість інакше: через навчання на історичних масивах даних, які вже обтяжені соціальними, гендерними чи расовими стереотипами минулих десятиліть. Якщо корпоративна база даних фіксує багаторічну штучну перевагу певного типу кандидатів при працевлаштуванні, математична предиктивна модель хибно сприймає цей дискримінаційний стереотип як об'єктивний критерій життєвого успіху та починає автоматично, на етапі скринінгу, відсіювати представників інших соціальних груп. Так звана «вбудована упередженість» стає невід'ємним токсичним ядром алгоритму, повністю руйнуючи саму сутність правової справедливості.

Подолання феномену алгоритмічної упередженості є об'єктивно неможливим лише через судові оскарження наслідків таких рішень постфактум. Воно вимагає створення превентивних, дієвих інституційних механізмів контролю на державному рівні. На наше глибоке переконання, практична імплементація цифрового сумління безальтернативно потребує створення незалежного спеціалізованого державного органу нагляду.

До виключних владних повноважень такої інституції повинно належати проведення обов'язкової процедури Legal & Algorithmic Impact Assessment (LAIA) – комплексної правової та техніко-алгоритмічної оцінки впливу до моменту введення будь-якої управлінської системи в експлуатацію. Ця процедура має нормативно передбачати жорстку державну сертифікацію «чистоти» та репрезентативності навчальних даних, стрес-тестування математичних моделей на виявлення прихованих дискримінаційних кореляцій та глибокий аудит відкритості програмного коду для всіх без винятку систем, що плануються до роботи в публічному секторі. Використання будь-якого несертифікованого державою алгоритму для ухвалення рішень, що впливають на соціальні чи економічні права громадян, має тягнути за собою невідворотну адміністративну та кримінальну відповідальність.

Надзвичайно важливим контекстом для розуміння меж імплементації захисних алгоритмічних механізмів є багата практика Європейського суду з прав людини (далі – ЄСПЛ), яка історично слугує еталоном та вектором для гармонізації українського національного законодавства.

Дослідник І. М. Берназюк ґрунтовно аналізує безпрецедентні виклики, які неконтрольований штучний інтелект ставить перед системою гарантій Європейської конвенції з прав людини [15, с. 91].

Право на повагу до приватного життя, беззаперечно гарантоване статтею 8 Конвенції про захист прав людини і основоположних свобод [16], сьогодні піддається колосальному системному тиску через розвиток технологій розпізнавання облич, обробки великих даних та впровадження предиктивної поліції (predictive policing).

Суттєвим правовим прецедентом у цій новітній сфері є нещодавнє рішення ЄСПЛ у справі «Glukhin v. Russia» (заява № 11519/20), в якому Суд прямо визнав використання технології розпізнавання облич (Facial Recognition Technologies) на мирній акції протесту без належного правового регулювання, запобіжників та незалежного інституційного контролю «високоінтрузивним» втручанням у приватність, що становить грубе порушення фундаментальних прав людини [17].

Не менш показовим та доктринально значущим є рішення окружного суду Гааги (Нідерланди) у гучній справі SyRI (System Risk Indication), де застосування урядом автоматизованої системи алгоритмічного виявлення ризиків шахрайства у чутливій сфері соціального забезпечення було визнано таким, що порушує межі приватності та статтю 8 ЄКПЛ через абсолютний брак прозорості алгоритмів оцінки та порушення принципу пропорційності збору персональних даних уразливих верств населення [18].

Зазначені міжнародні судові прецеденти формують чіткий і непорушний європейський стандарт: жодна, навіть найвища, декларована ефективність державного управління чи фіскального контролю не може слугувати виправданням для перетворення вільних громадян на безправних об'єктів тотального алгоритмічного нагляду. Права людини не можуть бути принесені в жертву технологічній зручності бюрократичного апарату.

Розглядаючи виклики цифровізації, неможливо оминати сферу захисту свободи слова. Як наголошує О. О. Карапетян, одним із найактуальніших викликів для адміністративно-правового забезпечення свободи вираження поглядів є стрімка поява генеративного ШІ та технологій deepfake, які здатні масово генерувати фейкові новини, маніпулятивні наративи та дезінформацію, підбиваючи довіру до інституцій [19, с. 93].

У відповідь на ці загрози держава часто вдається до автоматизованої онлайн-цензури. Проте неврегульованість механізмів алгоритмічної модерації контенту створює так званий «chilling effect» (ефект стримування), коли під обмеження можуть потрапляти цілком законні публікації, що висловлюють критичну оцінку дій влади. Відтак, механізми імплементації цифрового сумління у сфері медіа повинні виключати використання систем ШІ для непропорційного та непрозорого цензурування, вимагаючи обов'язкового маркування згенерованого контенту та збереження права особи на оскарження дій модераційних алгоритмів соціальних мереж.

Для сучасної України, яка щоденно перебуває у стані відбиття повномасштабної, жорстокої збройної агресії та функціонує в умовах правового режиму воєнного стану, питання регулювання та використання штучного інтелекту набувають вкрай гострого безпекового та екзистенційного забарвлення. Збройні Сили України, розвідка та весь сектор оборони активно та успішно впроваджують інтелектуальні автономні системи для швидкого аналізу розвідувальних даних, точного управління ударними дронами та математичного моделювання багаторівневих бойових сценаріїв. Проте, як дуже слушно і філософськи глибоко наголошує І. П. Марків, складний алгоритм машинного навчання дійсно може врахувати ймовірність тактичних дій противника, але він ніколи, за жодної обчислювальної потужності, не зможе осягнути аксіологічних та глибинних мотивів людської поведінки на полі бою – емоцій, екзистенційного страху, самопожертви, традицій чи морального обов'язку [20, с. 101].

Війна не є чистою математичною чи шаховою моделлю, і штучний інтелект не володіє розумінням вищої цінності людського життя, яке імперативно закріплене нормами міжнародного гуманітарного права (Женевськими конвенціями).

Відтак, впровадження своєрідного «аксіологічного фільтра» у військові та цивільні системи безпеки є питанням фізичного збереження людяності та недопущення воєнних злочинів в умовах технологічної війни. Це імперативно вимагає безапеляційного закріплення на найвищому законодавчому рівні принципу значущого людського контролю – «Human-in-the-loop» (людина в контурі управління). Машина може збирати терабайти даних, агрегувати інформацію та ранжувати потенційні цілі чи загрози, але фінальне рішення про застосування летальної сили або ухвалення доленосних адміністративних рішень щодо життєдіяльності та долі громадян повинно безальтернативно належати виключно людині, яка наділена правосвідомістю і несе за свій наказ повну, персональну юридичну (в тому числі міжнародно-правову) відповідальність.

Трансформація суспільних відносин під впливом алгоритмізації не обмежується питаннями публічного адміністрування чи військової безпеки; вона фундаментально перекроє економічний ландшафт та систему соціально-трудоу гарантій. Сучасні реалії диктують нагальну потребу перегляду

класичної парадигми відносин «держава – людина» у сфері соціального захисту. Інтеграція штучного інтелекту в економічні процеси призводить до виникнення безпрецедентного виклику – масового розмивання поняття традиційних трудових відносин внаслідок поширення так званої «гіг-економіки» та платформної зайнятості. Алгоритмічне управління на цифрових платформах (наприклад, службах доставки чи пасажирських перевезень) фактично виконує всі функції класичного менеджменту: розподіляє завдання, встановлює тарифи, контролює якість через системи рейтингів та застосовує санкції аж до алгоритмічного блокування акаунту (деплатформізації), що дорівнює звільненню.

Проте, формально-юридично такі платформи уникають статусу роботодавця, визначаючи себе лише як «інформаційних посередників». Це позбавляє мільйони економічно активних громадян базового доступу до класичних форм соціального захисту: оплачуваних лікарняних, гарантованих відпусток, захисту від травм на виробництві та повноцінного пенсійного страхування. Така правова трансформація вимагає негайного законодавчого визнання цифрової платформи податковим агентом та запровадження нових фіскальних інструментів для подолання кризи наповнення соціальних фондів [14, с. 111].

Крім економічної незахищеності, впровадження ШІ у сферу державних соціальних послуг створює ризики технологічного відчуження та порушення принципу рівного доступу. Держава, прагнучи оптимізувати бюджетні витрати, все частіше делегує функції нарахування пенсій, верифікації субсидій чи призначення соціальної допомоги автоматизованим системам оцінки ризиків. І тут знову виникає гостра потреба у впровадженні «аксіологічного фільтра» та імперативу «Human-in-the-loop». Алгоритм, навчений на історичних даних, може містити приховані упередження, які призводять до автоматичної і масової відмови у наданні допомоги найбільш вразливим верствам населення за непрямыми дискримінаційними ознаками (наприклад, за місцем реєстрації, що корелює з певним соціальним статусом).

Відповідно, у системі соціального захисту цифрової епохи має бути законодавче закріплено принцип: жодне автоматизоване рішення (Automated Decision-Making), що обмежує соціально-економічні права громадянина, відмовляє йому у наданні життєво необхідних послуг або зменшує їх обсяг, не може набувати юридичної сили без обов'язкової верифікації живою посадовою особою та без надання особі чіткого пояснення логіки такого рішення (Explainable AI). Ситуація, коли державний орган мотивує відмову у виплаті аргументом «так вирішила система», є неприпустимою ознакою алгоритмічного свавілля, що руйнує довіру до інститутів влади та позбавляє особу реального права на апеляційне оскарження такого рішення у судовому порядку.

Зіштовхуючись із викликом технологічного безробіття та знецінення традиційних людських навичок перед обличчям генеративного штучного інтелекту, доктрина прав людини має еволюціонувати. Пасивні методи соціального захисту, такі як виплата допомоги по безробіттю або навіть дискусійне запровадження безумовного базового доходу, не вирішують фундаментальної проблеми – екзистенційної ізоляції людини від суспільно-корисної праці та втрати нею економічної суб'єктності.

У зв'язку з цим, видається концептуально обґрунтованим підхід щодо необхідності формування та визнання нового суб'єктивного права людини четвертого покоління – права на безперервну технологічну адаптацію (Lifelong Reskilling Right) [14, с. 112]. Це право передбачає обов'язок держави та великих технологічних корпорацій не просто компенсувати втрату доходу, а створити інституційні умови (активний «соціальний ліфт») для постійного безкоштовного перенавчання громадян, гарантуючи їм доступ до нових компетенцій, які дозволять ефективно взаємодіяти з системами ШІ, а не бути витисненими ними на маргінес суспільства.

Впровадження архітектури цифрового сумління у соціальну та правову сфери вимагає системної розбудови інститутів алгоритмічної підзвітності. Це передбачає не лише встановлення технічних стандартів для розробників програмного забезпечення, але й докорінну зміну процесуального законодавства. Зокрема, у спорах щодо дискримінації або незаконного позбавлення соціальних прав через дії автоматизованої системи має бути запроваджена презумпція «алгоритмічної вини». Це означає, що тягар доведення коректності, неупередженості та законності роботи алгоритму покладається виключно на власника або оператора такої системи (державний орган чи корпорацію), а не на громадянина-позивача, який априорі позбавлений доступу до вихідного коду та баз навчальних даних цієї системи. Лише за умови такого перерозподілу процесуального тягаря можна забезпечити реальну рівність сторін у цифровому правосудді.

Підсумовуючи проведене науково-теоретичне дослідження, варто констатувати, що концепція «архітектури цифрового сумління» не є абстрактною філософською метафорою, а постає безальтернативним механізмом збереження антропоцентризму права в умовах технологічної сингулярності.

Стрімка алгоритмізація публічного управління, правосуддя, соціального захисту та військової сфери вимагає від держави проактивних дій. Трансформація етичних декларацій у жорсткий юридичний код, запровадження інститутів обов'язкового алгоритмічного аудиту, створення спеціалізованих органів нагляду за ШІ, адаптація соціально-економічного законодавства до реалій платформної економіки та, найголовніше, імперативне збереження змістовного людського контролю над усіма життєво важливими процесами (Human-in-the-loop) є необхідними умовами виживання правової держави. Україна, яка сьогодні формує передовий досвід цифровізації державних послуг в екстремальних умовах війни, має унікальний історичний шанс стати не лише споживачем глобальних технологій, але й творцем нових еталонних стандартів цифрової юриспруденції, де технологічний прогрес завжди залишатиметься підпорядкованим гідності, свободі та правам людини.

**Висновки.** Підсумовуючи вищенаведене, необхідно зазначити, що архітектура цифрового сумління постає не просто як технічне завдання, а як фундаментальна умова виживання правової держави в епоху технологічної сингулярності. Доведено, що лише через інтеграцію правових імперативів безпосередньо у програмний код за принципом Law-by-Design можна гарантувати непорушність антропоцентризму права. Цифрове сумління має стати жорстким алгоритмічним фільтром, що унеможливило прийняття будь-яких дискримінаційних чи антигуманних рішень автоматизованими системами.

Перспективи подальших розвідок у даному напрямку вбачаються у детальному дослідженні правового регулювання генеративного штучного інтелекту та розробці галузевих стандартів цифрового сумління для сфер правосуддя, медицини та оборони. Окремої уваги потребує питання міжнародної координації стандартів алгоритмічної підзвітності, оскільки транскордонна природа цифрових технологій робить національне регулювання ефективним лише за умови його гармонізації з глобальними гуманістичними нормами. Україна має потенціал стати лідером у формуванні етичного коду ШІ, що базується на унікальному досвіді захисту людських цінностей у часи глобальних викликів.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Подковенко Т. О. Принцип антропоцентризму в контексті сучасного праворозуміння. *Юридичний вісник*. 2012. № 2(23). С. 22–25.
2. Конституція України: Закон України від 28.06.1996 № 254к/96-ВР. Відомості Верховної Ради України. 1996. № 30. Ст. 141. URL: <https://surl.li/mfkokr> (дата звернення: 24.03.2026).
3. Пашко Л. А. Щодо обов'язковості використання пріоритетних запобіжників дегуманізації цифровізації Української держави. *Вчені записки ТНУ імені В. І. Вернадського*. Серія: Публічне управління та адміністрування. 2025. Том 36 (75) № 3. С. 122–127. URL: <https://surl.li/ycydtu> (дата звернення: 24.03.2026).
4. Коломоєць І. Цифровий гуманізм як соціокультурна парадигма гуманістично-орієнтованого суспільства цифрової епохи. *Humanities Studies*. 2025. Вип. 25 (102). С. 56–67. URL: <https://surl.li/qwfkxp> (дата звернення: 25.03.2026).
5. Пікуля Т. О. Трансформація політико-правових доктрин держави і права під впливом розвитку штучного інтелекту: європейська концепція врядування штучним інтелектом. *Актуальні проблеми права: теорія і практика*. 2026. № 1 (51). С. 180–189. URL: <https://journals.snu.edu.ua/index.php/app/article/view/1233>
6. Сікало М. В. Від фрагментації до конвергенції: теоретична архітектура транзитивної державності у цифрову епоху. *Державне будівництво*. 2025. № 1 (37). С. 152–170. URL: <https://surl.li/fcewei> (дата звернення: 26.03.2026).
7. Міщенко В. І. Механізми регулювання процесів цифрування для забезпечення національної стійкості економічного розвитку. *Економічний простір*. 2024. № 189. С. 284–289. URL: <https://surl.li/zqobel> (дата звернення: 26.03.2026).
8. Кравчук С., Василенко С. Суб'єктивні права людини в цифрову епоху. *Вісник Національного університету “Львівська політехніка”*. Серія: “Юридичні науки”. 2025. № 2 (46). С. 146–155. URL: <https://surl.li/etzgom> (дата звернення: 26.03.2026).

9. General Data Protection Regulation 2016/679. OJ L 119, 04.05.2016. URL: <https://gdpr-info.eu/> (дата звернення: 26.03.2026).
10. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. Official Journal of the European Union, L 2024/1689. URL: <https://eurlex.europa.eu/eli/reg/2024/1689/oj/eng> (дата звернення: 26.03.2026).
11. Сарнацький М. Штучний інтелект як новий суб'єкт адміністративної взаємодії: виклики правової легітимності в контексті державних закупівель. *Dictum factum*. 2025. № 2 (18). С. 129–137. URL: <https://surl.lt/umurer> (дата звернення: 26.03.2026).
12. Про публічні закупівлі: Закон України від 25.12.2015 № 922-VIII. Відомості Верховної Ради України. 2016. № 9. Ст. 89. URL: <https://zakon.rada.gov.ua/laws/show/922-19> (дата звернення: 26.03.2026).
13. Сарнацький М. Використання штучного інтелекту в електронних системах публічних закупівель: правові аспекти, виклики та перспективи. *Dictum factum*. 2025. № 1 (17). С. 78–86. URL: <https://surl.li/bnhdha> (дата звернення: 26.03.2026).
14. Кравчук С. М. Нові форми соціального захисту людини в умовах розвитку штучного інтелекту. *Науковий вісник Ужгородського національного університету*. Серія: Право. 2026. Вип. 93. Ч. 1. С. 107–113. URL: <https://surl.li/dzucsq> (дата звернення: 26.03.2026).
15. Берназюк І. М. Штучний інтелект і права людини: виклики для Європейської конвенції з прав людини. *Аналітично-порівняльне правознавство*. 2025. № 1(3). С. 89–99. URL: <https://surl.li/bhofot> (дата звернення: 26.03.2026).
16. Конвенція про захист прав людини і основоположних свобод від 04.11.1950. URL: <https://surl.li/zvjhuo> (дата звернення: 26.03.2026).
17. Case of Glukhin v. Russia: decision of the European Court of Human Rights, 04.07.2023, № 11519/20. URL: <https://surl.lt/vhtcqq> (дата звернення: 26.03.2026).
18. Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case. Oxford Academic. URL: <https://surl.li/lvobti> (дата звернення: 26.03.2026).
19. Карапетян О. О. Адміністративно-правові аспекти захисту свободи слова в умовах цифрової трансформації медіа. 2025. URL: <https://surl.lu/qzzrtg> (дата звернення: 26.03.2026).
20. Марків І. П. Використання алгоритмів штучного інтелекту в процесі науково-інформаційного осмислення військових кейсів: між аналізом і сенсом. *Вісник гуманітарних наук*. 2025. № 8. С. 100–104. URL: <https://surl.li/fmipaj> (дата звернення: 26.03.2026).

Дата першого надходження рукопису до видання: 26.03.2026

Дата прийняття до друку рукопису після рецензування: 23.04.2026

Дата публікації: 10.05.2026

© Кравчук С.М., Коваль А.В., 2026

Стаття поширюється на умовах ліцензії CC BY 4.0