

Страсбург, 19 грудня 2025 року

CEPEJ(2025)18Final

ЄВРОПЕЙСЬКА КОМІСІЯ З ПИТАНЬ ЕФЕКТИВНОСТІ ПРАВОСУДДЯ (СЕРЕЈ)

КЕРІВНІ ПРИНЦИПИ ЩОДО ВИКОРИСТАННЯ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ В СУДАХ

*Ухвалено на 45-му пленарному засіданні CEPEJ
(Страсбург, 4–5 грудня 2025 року).*

Зміст

ВИКЛАД ОСНОВНИХ ПОЛОЖЕНЬ	3
1. ВСТУП	5
1.1. Преамбула	5
1.2. Генеративний штучний інтелект	5
1.3. Генеративний штучний інтелект у сфері правосуддя	7
2. КЕРІВНІ ПРИНЦИПИ	9
2.1. Операційні аспекти використання генеративного штучного інтелекту в судах	9
2.1.1. Визначення проблеми має передувати вибору технічного рішення	9
2.1.2.1. Управління справами та судово адміністрування	10
2.1.2.2. Управління документами та обробка інформації	11
2.1.2.3. Інформаційні сервіси та допомога користувачам	11
2.1.2.4. Підтримка складання документів та судова документація	11
2.1.3. Урахування неприйнятних сфер застосування	12
2.1.4. Надання переваги індивідуалізованим і безпечним інструментам порівняно з універсальними готовими рішеннями	13
2.1.5. Забезпечення належного управління під час упровадження систем генеративного штучного інтелекту	16
2.1.6. Усвідомлення того, як сторони використовують генеративний штучний інтелект	19
2.1.7. Забезпечення обізнаності та навчання користувачів штучного інтелекту	20
2.2. Нормативні аспекти використання генеративного штучного інтелекту в судах	21
2.2.1.1. Гарантії справедливого судового розгляду	22
2.2.1.2. Принцип законності, правової визначеності та передбачуваності	22
2.2.1.3. Виключність судової влади	23
2.2.1.4. Обов'язкова відповідальність людини та нагляд	23
2.2.1.5. Суддівська незалежність і необов'язковий характер використання штучного інтелекту	23
2.2.1.6. Прозорість та належне обґрунтування	24
2.2.1.7. Право на доступ до суду	24
2.2.2. Право на свободу та особисту недоторканність (стаття 5 ЄКПЛ)	25
2.2.3. Право на повагу до приватного життя та захист персональних даних (стаття 8 ЄКПЛ)	25
2.2.4. Право на ефективний засіб правового захисту (стаття 13 ЄКПЛ)	26
2.2.5. Заборона дискримінації (стаття 14 ЄКПЛ та Протокол № 12)	26
2.2.5.1. Принцип недискримінації	26
2.2.5.2. Нейтральність та активне запобігання упередженості	27
2.2.6. Додаткові запобіжники	27
2.2.6.1. Пояснюваність та прозоре обґрунтування	27
2.2.6.2. Пропорційність і обережність у впровадженні генеративного штучного інтелекту	28
2.2.6.3. Відповідальність держави за шкоду, завдану використанням генеративного штучного інтелекту	28
2.2.6.4. Принцип процесуальної справедливості та ефективності	28
3. ГЛОСАРІЙ	29

ВИКЛАД ОСНОВНИХ ПОЛОЖЕНЬ

Ці Керівні принципи визначають засади впровадження генеративного штучного інтелекту (ШІ) у сфері здійснення правосуддя з акцентом на забезпеченні правової визначеності, суддівської незалежності за умови ефективного людського нагляду, прозорості та простежуваності.

На початковому етапі універсальні інструменти, розроблені великими технологічними компаніями, можуть розглядатися лише як перший крок; надалі доцільно переходити до спеціалізованих рішень і, за потреби, до локальних великих мовних моделей (LLM) або їх розгортання із суверенним контролем над даними та інфраструктурою. Вкрай важливо, щоб такі інструменти були адаптовані до судових процесів і впроваджувалися під публічним контролем.

Впровадження таких рішень має здійснюватися з дотриманням Європейської конвенції з прав людини (ЄКПЛ), Рамкової конвенції Ради Європи про штучний інтелект (CETS № 225), Конвенції 108+ про захист осіб у зв'язку з обробкою персональних даних, а також Європейської етичної хартії CEPEJ щодо використання штучного інтелекту в судових системах та їх середовищі.

Ці Керівні принципи ґрунтуються на таких ключових принципах: здійснення судової влади є виключною компетенцією судів; ефективний доступ до судді гарантується завжди; результати, згенеровані генеративним штучним інтелектом, ніколи не є обов'язковими; використання генеративного штучного інтелекту в судовій діяльності та під час підготовки судових рішень має бути прозорим. Вимоги щодо захисту приватності вимагають псевдонімізації даних, а також забезпечення виключного контролю над даними та інфраструктурою з боку публічного органу, відповідального за надання інформаційно-технологічних послуг судовій владі. Вирішального значення набуває дотримання законності, суддівської незалежності, недискримінації, права на ефективний засіб юридичного захисту, а також простежуваності й пояснюваності обґрунтування.

Запропоновано поетапний підхід до впровадження, що передбачає пілотні проєкти, визначення показників та можливість перегляду обраного підходу за потреби. Такий підхід ґрунтується на принципі субсидіарності, який наголошує на визначенні функціональних потреб і пріоритетному розгляді традиційних технологічних рішень. Генеративний штучний інтелект слід застосовувати лише у випадках, коли альтернативні рішення визнаються неефективними або нераціональними.

Використання генеративного штучного інтелекту може бути корисним у різних професійних контекстах, зокрема в управлінні справами, обробці документів, наданні публічної інформації та підтримці підготовки типових судових документів. У разі його застосування для надання допомоги судовим органам у дослідженні та тлумаченні фактичних обставин і права, у застосуванні права до конкретних сукупностей фактів або в будь-якій іншій формі підтримки судочинства, мають дотримуватися суворі вимоги. Для забезпечення належного впровадження необхідним є проведення ґрунтовної попередньої (ex ante) оцінки ризиків і впливу до будь-якого розгортання таких систем. Запровадження механізмів безперервного моніторингу та зворотного зв'язку дає змогу судам і системі правосуддя загалом своєчасно вносити коригування та вдосконалення. Нарешті, підготовка всіх працівників судів і користувачів системи є ключовою умовою для забезпечення її ефективного використання.

Держава має залишатися відповідальною за будь-яку шкоду, завдану внаслідок використання систем генеративного штучного інтелекту у сфері здійснення правосуддя, гарантуючи користувачам ефективні шляхи звернення та засоби правового захисту у разі завдання шкоди.

1. ВСТУП

1.1. Преамбула

1. Ці Керівні принципи мають на меті надати практичні рекомендації щодо безпечного впровадження генеративного штучного інтелекту (ШІ) у сфері здійснення правосуддя. Вони адресовані органам, відповідальним за здійснення правосуддя, а також суддям і працівникам судової системи. Керівні принципи підлягають періодичному перегляду та оновленню з урахуванням новітніх технологічних розробок.

2. Використання штучного інтелекту в системі правосуддя має здійснюватися з дотриманням прав, гарантованих Європейською конвенцією з прав людини (ЄКПЛ), Рамковою конвенцією Ради Європи про штучний інтелект і права людини, демократію та верховенство права (CETS № 225), а також вимог щодо захисту персональних даних, установлених Конвенцією Ради Європи про захист персональних даних (CETS № 108), принципів, викладених у Європейській етичній хартії щодо використання штучного інтелекту в судових системах та їх середовищі (Хартія CEPEJ), і, за наявності застосовності, Регламенту (ЄС) 2024/1689 Європейського Парламенту та Ради від 13 червня 2024 року, яким установлюються гармонізовані правила щодо штучного інтелекту.

3. Ці Керівні принципи були підготовлені Робочою групою CEPEJ з питань кіберправосуддя та штучного інтелекту (CEPEJ-GT-CYBERJUST) на основі внеску експертів.¹ Вони враховують попередні напрацювання CEPEJ, зокрема «Інформаційну записку щодо використання генеративного штучного інтелекту суддями та працівниками судової системи в робочому контексті», «Міркування Консультативної ради CEPEJ з питань штучного інтелекту щодо того, як забезпечити ефективне використання штучного інтелекту у сфері здійснення правосуддя», а також «Першу доповідь щодо використання штучного інтелекту (ШІ) у сфері здійснення правосуддя», підготовлену на основі інформації, що міститься в Ресурсному центрі CEPEJ з питань кіберправосуддя та штучного інтелекту. Керівні принципи також ілюструються прикладами інструментів, представлених у Ресурсному центрі CEPEJ з питань кіберправосуддя та штучного інтелекту.

1.2. Генеративний штучний інтелект

4. Генеративний штучний інтелект є напрямом штучного інтелекту, що

¹ Хав'єр Ерсілья Гарсія, суддя, Іспанія.

зосереджується на навчанні машин створювати новий контент, який переконливо імітує результати діяльності людини. На відміну від систем, що лише класифікують або прогнозують дані, такі системи навчаються на великих масивах даних для виявлення базових закономірностей і структур, що дає їм змогу генерувати текст, зображення, музику, відео та інші види контенту.²

5. У сфері генерації тексту розвиток інструментів обробки природної мови (natural language processing) дав змогу автоматизованим програмним засобам коректно інтерпретувати та генерувати мовлення. Раніше розуміння людської мови становило складне завдання для систем штучного інтелекту через її багатозначність і складність, однак завдяки досягненням у сфері глибокого навчання моделі стали здатними виявляти закономірності мови та створювати зв'язні відповіді або тексти. Це зумовило суттєву зміну підходів до обробки та створення машинами контенту на основі текстових даних.

6. Технологія, що лежить в основі генеративного штучного інтелекту, ґрунтується на глибокому навчанні та сучасних архітектурах нейронних мереж. По суті, ці моделі, відомі як великі мовні моделі (LLM), навчаються на значних обсягах даних для виявлення складних закономірностей (зокрема послідовностей слів, пікселів зображень або музичних нот), що дає змогу генерувати новий контент, який відтворює характеристики вихідного матеріалу.

7. Незважаючи на ефективність у виконанні широкого спектра завдань і поширене застосування в чат-ботах, генеративний штучний інтелект на поточному етапі розвитку має притаманні йому обмеження, одним із проявів яких є так звані «галюцинації». У сфері генерації природної мови «галюцинаціями» є результати, які модель створює без належного фактичного підґрунтя у вхідних даних або довідкових знаннях: вони можуть прямо суперечити джерельному тексту (внутрішні) або містити неможливі чи хибні дані (зовнішні). По суті, йдеться про вигадану інформацію, яку система подає як факт.³

8. Інші суттєві обмеження стосуються потенційних питань авторського права, пов'язаних із навчальними даними, використаними для первинного навчання таких моделей, а також ризику ненавмисного розкриття вхідних даних третім особам і їх подальшого використання для навчання генеративних моделей штучного інтелекту, що може призводити до порушення застосовного

² Варанкар, М., та Патіл, Р. (2024). Генеративний штучний інтелект. International Journal of Scientific Research in Engineering and Management, 8(04). [doi:10.55041/IJSREM31146](https://doi.org/10.55041/IJSREM31146)

³ Джі, Ч., Лі, Н., Фріске, Р., Ю, Т., Су, Д., Сю, Ю., Ісїї, Е., Банґ, Ю., Чен, Д., Дай, В., Чан, Х. С., Мадотто, А., та Фанґ, П. (2022). Огляд проблеми «галюцинацій» у генерації природної мови. arXiv. <https://arxiv.org/abs/2202.03629>

законодавства у сфері захисту персональних даних.

1.3. Генеративний штучний інтелект у сфері правосуддя

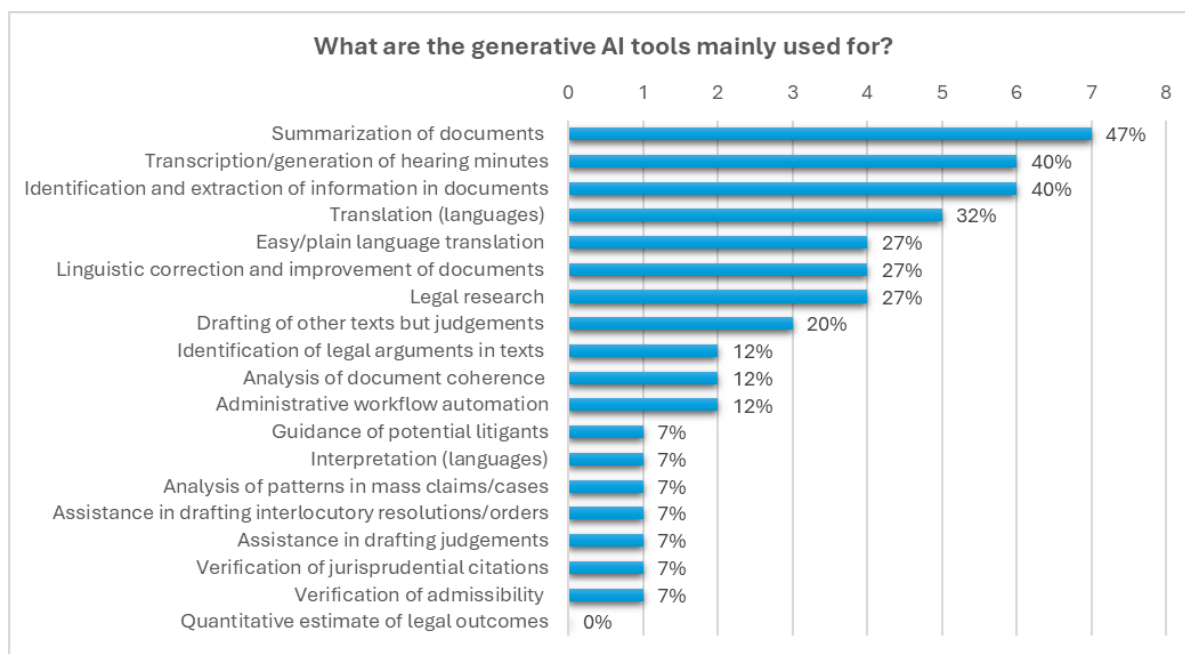
9. Використання рішень генеративного штучного інтелекту правниками дедалі зростає, зокрема для створення стислих викладів, редагування текстів, формування правових позицій на основі судової практики, а також для взаємодії зі сторонами у справі (зокрема з метою пояснення судових рішень клієнтам) тощо. Згідно зі звітом *Future Ready Lawyer 2024*,⁴ у межах якого було опитано 712 правників у Сполучених Штатах Америки та дев'яти європейських країнах, 68 % юридичних фірм і 76 % корпоративних юридичних підрозділів використовують генеративний штучний інтелект щонайменше раз на тиждень, при цьому понад третина застосовує його щоденно (35 % – у корпоративних юридичних підрозділах і 33 % – у юридичних фірмах). У Європі найвищий рівень щоденного використання генеративного штучного інтелекту зафіксовано в Нідерландах (42 %), за якими йде Німеччина (38 %). Основними напрямками застосування таких інструментів є автоматизація підготовки документів (67 %), управління часом (66 %) та стратегічне планування (62 %). Крім того, 58 % юридичних фірм і 73 % корпоративних юридичних підрозділів планують збільшити інвестиції в штучний інтелект протягом наступних трьох років.

10. Згідно з опитуванням членів Європейської мережі CEREJ з кіберправосуддя, проведеним у жовтні 2025 року, 46 % респондентів⁵ підтвердили використання генеративного штучного інтелекту в судах. При цьому 74 % із них зазначили застосування готових (універсальних) рішень (зокрема Microsoft Copilot, ChatGPT, DeepL), а 52 % – індивідуалізованих (адаптованих) рішень (можливий вибір обох варіантів). Найчастіше генеративний штучний інтелект використовується для узагальнення документів та виконання редакторських і допоміжних функцій. Жоден із респондентів не повідомив про його використання для кількісної оцінки правових результатів (див. рисунок 1).

⁴ Wolters Kluwer. (2024). Звіт «Future Ready Lawyer 2024»: правничі інновації – скористатися майбутнім чи відстати? Wolters Kluwer Legal & Regulatory. <https://www.wolterskluwer.com/en/know/future-ready-lawyer-2024>

⁵ Опитування охопило 33 респондентів із 43 держав – членів Ради Європи, до яких було звернено запит. Показники заокруглено до цілих чисел.

Рисунок 1. Для яких цілей переважно використовуються інструменти генеративного штучного інтелекту?



Узагальнення документів – 47 %

Транскрибування та підготовка протоколів судових засідань – 40 %

Виявлення та витяг інформації з документів – 40 %

Переклад (мови) – 32 %

Переклад простою (зрозумілою) мовою – 27 %

Лінгвістичне редагування та вдосконалення документів – 27 %

Правове дослідження (legal research) – 27 %

Підготовка інших текстів, окрім судових рішень – 20 %

Виявлення правових доводів у текстах – 12 %

Аналіз логічної зв'язності документів – 12 %

Автоматизація адміністративних робочих процесів – 12 %

Надання орієнтовних роз'яснень потенційним сторонам у справі – 7 %

Усний переклад (мови) – 7 %

Аналіз закономірностей у масових позовах / справах – 7 %

Допомога в підготовці проміжних судових рішень / ухвал – 7 %

Допомога в підготовці судових рішень – 7 %

Перевірка посилань на судову практику – 7 %

Перевірка допустимості звернень – 7 %

Кількісна оцінка правових результатів – 0 %

Рисунок 1. Використання генеративного штучного інтелекту в судах.

11. У різних державах використання інструментів генеративного штучного інтелекту в органах публічної адміністрації та судах супроводжується укладенням угод про конфіденційність, захистом даних про використання, а також наданням гарантій того, що такі дані не будуть повторно використані для навчання моделей.^{6 7}

12. У правничій сфері використання генеративного штучного інтелекту вже призводило до судових проваджень, у межах яких подавалися процесуальні документи з хибними посиланнями на судову практику,^{8 9} що в окремих випадках призводило до застосування санкцій до користувачів таких інструментів.¹⁰ Оскільки «галюцинації» відповідно до сучасного рівня розвитку технологій є неминучими¹¹ або, у найкращому разі можуть бути лише пом'якшені,¹² це становить ризик, який слід враховувати під час використання генеративного штучного інтелекту у сфері правосуддя за відсутності належного контролю з боку людини.

2. КЕРІВНІ ПРИНЦИПИ

2.1. Операційні аспекти використання генеративного штучного інтелекту в судах

2.1.1. Визначення проблеми має передувати вибору технічного рішення

13. Принцип технологічної нейтральності передбачає, що органи публічної влади мають зосереджуватися на функціональних потребах і, виходячи з них, обирати технологічні рішення, здатні адаптуватися з часом. Такий підхід спрямований на мінімізацію технологічної залежності, уникнення нав'язування конкретних технічних реалізацій або продуктів та збереження гнучкості в

⁶ Artificial Lawyer. (2025, 24 квітня). Суди Великої Британії розгортають Microsoft Copilot для суддів та оновлюють правила щодо генеративного ШІ. <https://www.artificiallawyer.com/2025/04/24/uk-courts-roll-out-microsoft-copilot-for-judges-update-genai-rules/>

⁷ El País. (2025, 22 вересня). Чи знаєте ви вже, що таке carpeta.justicia.es? Отримано з <https://cincodias.elpais.com/extras/2025-09-22/ya-sabes-lo-que-es-carpetajusticiaes.html>

⁸ Вайзер, Б. (2023, 27 травня). Ось що відбувається, коли ваш адвокат використовує ChatGPT. The New York Times. <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>

⁹ MIT Sloan Teaching & Learning Technologies. (2024, 12 листопада). Коли ШІ помиляється: подолання галюцинацій і упередженості штучного інтелекту. <https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/>

¹⁰ Каррік, Д., та Кестевен, С. (2023, 24 червня). Цей адвокат у США використав ChatGPT для підготовки правової позиції – з прикрими наслідками. Усі ми можемо винести урок із цієї помилки. ABC News. <https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068>

¹¹ Сю, Ч., Джейн, С., & Канканхаллі, М. (2024). Галюцинації є неминучими: вроджене обмеження великих мовних моделей. arXiv. <https://arxiv.org/abs/2401.11817>

¹² Ніє, Ф., Яо, Дж.-Г., Ван, Дж., Пан, Ж., & Лін, Ч.-Й. (2019, липень). Простий підхід до зменшення галюцинацій у нейронній генерації поверхневих структур. У матеріалах 57-го щорічного засідання Асоціації з обчислювальної лінгвістики, с. 2673–2679. Асоціація з обчислювальної лінгвістики.

умовах стрімкого технологічного розвитку. Цей принцип має застосовуватися і до використання генеративного штучного інтелекту у сфері судового адміністрування через принцип субсидіарності, відповідно до якого судові адміністрації, по-перше, чітко визначають функціональні потреби, а по-друге, розглядають різні технологічні рішення – як традиційні, так і інноваційні, – перш ніж обрати найбільш придатне та найменш ризиковане. Такий підхід не повинен перешкоджати судам у прагненні підвищувати ефективність здійснення правосуддя із застосуванням технологій там, де це є доцільним, а також у прийнятті рішення не впроваджувати генеративний штучний інтелект в усталених випадках використання, які вже відповідають чітко визначеним функціональним потребам.

14. Як суди, так і органи судового адміністрування стикаються з низкою викликів, зокрема нестачею персоналу, великими обсягами масових позовів, повторюваними завданнями з підготовки документів, неефективним управлінням матеріалами справ, процесуальними затримками, труднощами у пошуку судової практики, проблемами комунікації з користувачами системи правосуддя, а також відсутністю інтероперабельності між чинними інформаційно-технологічними системами.

15. Після визначення функціональних потреб і проблем у межах судового адміністрування наступним кроком є відбір відповідних технічних рішень відповідно до принципів технологічної нейтральності та субсидіарності у використанні штучного інтелекту.

16. Під час відбору найбільш придатних технічних рішень слід враховувати їхню економічну доцільність, зокрема такі чинники, як споживання енергії, заходи кібербезпеки, умови ліцензування, витрати на технічне обслуговування, розробку та консультаційні послуги, а також не обмежуватися оцінкою лише початкової вартості придбання.

2.1.2. Визначення сфер, придатних для використання генеративного штучного інтелекту¹³

2.1.2.1. Управління справами та судове адміністрування

17. Генеративний штучний інтелект може застосовуватися для прогнозування та управління робочим навантаженням шляхом аналізу історичних закономірностей, а також для створення сховищ судової практики, упорядкованих за видами справ і предметом спору. Такі інструменти можуть сприяти виявленню типових моделей аргументації у подібних позовах, що дає

¹³ Наведений перелік не є вичерпним.

зможу ідентифікувати провадження, які можуть бути вирішені уніфіковано у разі порушення однакового правового питання, а також підтримувати розподіл справ. Крім того, генеративний штучний інтелект може забезпечувати автоматизоване виявлення спірних питань шляхом порівняння позовних заяв і відзивів.

2.1.2.2. Управління документами та обробка інформації

18. Генеративний штучний інтелект може використовуватися для автоматизованого складання протоколів судових засідань, формування тематичних і хронологічних показників для значних за обсягом матеріалів справ, а також для витягу релевантних даних з неструктурованих документів. Крім того, він може застосовуватися для автоматичного виявлення ключової інформації, зокрема імен, дат і сум, із подальшим включенням таких відомостей до наперед визначених шаблонів.

2.1.2.3. Інформаційні сервіси та допомога користувачам

19. Генеративний штучний інтелект може забезпечувати створення спеціалізованих чат-ботів за окремими галузями права, а також інтерактивних систем процесуального орієнтування для громадян. Такі рішення мають розроблятися з особливою увагою до точності наданої інформації та зрозумілості комунікації з нефаховими користувачами, із застосуванням простої та доступної мови.

2.1.2.4. Підтримка складання документів та судова документація

20. Генеративний штучний інтелект може застосовуватися для автоматизованого створення стандартних процесуальних документів, надання допомоги у підготовці процесуальних ухвал, а також для перевірки внутрішньої узгодженості документів. Він може формувати структуровані узагальнення судових рішень, процесуальних документів і апеляційних скарг за умови обов'язкової перевірки суддею відповідності такого узагальнення оригіналу. Додатково, генеративний штучний інтелект може використовуватися для підготовки допоміжних матеріалів, зокрема звітів, планів або інтелектуальних карт, що надають узагальнений огляд питань, щодо яких існує згода або розбіжності.

21. Крім того, генеративний штучний інтелект може застосовуватися для аналізу позовних заяв і відзивів із формуванням пропозицій щодо релевантної судової практики, виявлення та узагальнення судової практики, що має

значення для кожної справи, попереднього аналізу апеляційних скарг щодо дотримання формальних і технічних вимог, а також для встановлення логічних зв'язків між доводами, викладеними в різних процесуальних документах, із формуванням структурованих матеріалів, у яких заперечення відповідача системно співвідносяться з кожним доводом позивача.

2.1.3. Урахування неприйнятних сфер застосування

22. Регламент ЄС про штучний інтелект (Регламент (ЄС) 2024/1689) установлює конкретні заборони на практики використання штучного інтелекту, які є особливо релевантними для сфери здійснення правосуддя, а також потенційно — для генеративного штучного інтелекту. Стаття 5 цього Регламенту передбачає такі заборонені практики використання ШІ, які можуть слугувати орієнтиром також для держав, на які не поширюється його дія: системи штучного інтелекту, що застосовують підсвідомі або маніпулятивні методи (стаття 5(1)(a)); системи штучного інтелекту, що експлуатують уразливості осіб або груп осіб (стаття 5(1)(b)); системи соціального рейтингування (стаття 5(1)(c)); системи оцінки ризиків для прогнозування вчинення кримінальних правопорушень (стаття 5(1)(d)); системи розширення баз даних розпізнавання облич (стаття 5(1)(e)); системи виведення емоцій (стаття 5(1)(f)).

23. Крім того, хоча такі системи не є прямо забороненими, Регламент ЄС про штучний інтелект відносить окремі системи ШІ у сфері правосуддя до категорії високого ризику відповідно до Додатка III, зокрема ті, що призначені для надання допомоги судовим органам у дослідженні та тлумаченні фактів і права, а також у застосуванні права до конкретних фактичних обставин. Такі системи підлягають дотриманню суворих вимог, однак не підлягають загальній забороні.

24. Хоча великі мовні моделі (LLM) можуть бути корисними на етапі збирання доказів у межах розслідування, вони не повинні замінювати судову оцінку доказів після їх долучення до судового розгляду. Незважаючи на існування окремих напрямів логічного та нейросимвольного штучного інтелекту, які інтегрують каузальні моделі,¹⁴ генеративний штучний інтелект не має здатності до «проспективної теоретизації та формування причинно-наслідкової логіки».^{15 16}

¹⁴ Коуллоу, Б. К., & Реджі, В. (2025). Нейросимвольний штучний інтелект у 2024 році: систематичний огляд. arXiv. <https://arxiv.org/abs/2501.05435>

¹⁵ Фелін, Т., та Голвер, М. (2024). Теорія — це все, що вам потрібно: штучний інтелект, людське пізнання та ухвалення рішень. Доступно за посиланням: <https://ssrn.com/abstract=4737265>.

¹⁶ Лі, С., Цай, Ц., Ван, С., Ю, К., & Чен, Ф. (2025). Огляд підходів до посилення здатності великих мовних моделей

25. Процес ухвалення рішень має залишатися під контролем людини. Справи, що потребують нового нормативного тлумачення або аналізу абстрактних правових принципів, не слід доручати штучному інтелекту, з огляду на те, що він не має *«внутрішнього розуміння того, що є істинним або хибним у реальності»*.¹⁷ Так само слід уникати використання генеративного штучного інтелекту у справах, які потребують еволюційного тлумачення нещодавно ухвалених нормативних актів, оскільки генеративний штучний інтелект може виявляти обмеження в адаптації свого розуміння до нового регулювання, якщо він не був спеціально навчений на відповідних даних.¹⁸

26. Аналіз конституційності або відповідності міжнародним договорам потребує рівня аксіологічного тлумачення, що перевищує нинішні можливості штучного інтелекту. Навіть якщо в майбутньому це було б інакше, таке питання може бути довірене виключно людському судженню, оскільки воно стосується способу, у який громадяни визначили порядок свого врядування. Відтак «дух часу» кожного періоду впливатиме на «контроль конституційності та конвенційності», і це не може бути делеговано на розсуд штучного інтелекту.

2.1.4. Надання переваги індивідуалізованим і безпечним інструментам порівняно з універсальними готовими рішеннями

27. Розмежування між універсальними комерційними інструментами генеративного штучного інтелекту та індивідуалізованими судовими інструментами генеративного штучного інтелекту має істотне значення для сфери здійснення правосуддя, оскільки воно стосується не лише технічних характеристик, а й фундаментальних питань захисту даних, судової незалежності та цілісності судового провадження.

28. Безоплатні комерційні інструменти генеративного штучного інтелекту становлять значні ризики. Коли судді та інші працівники суду використовують сервіси генеративного штучного інтелекту, що не є офіційно затвердженими, — як безоплатні, так і платні комерційні сервіси генеративного штучного інтелекту, — існує ризик ненавмисної міжнародної передачі даних, переважно до центрів обробки даних у Сполучених Штатах Америки, які перебувають у володінні компаній — постачальників таких сервісів. Надалі ці дані можуть використовуватися для повторного навчання моделей, що фактично

до каузального міркування. arXiv. <https://arxiv.org/abs/2503.09326>

¹⁷ Там само.

¹⁸ Ерсілья Гарсія, Х. (2025). Автоматизоване правосуддя: між штучними інтелектами, що імітують, і тими, що переконують. Lex Social, Revista de Derechos Sociales, 15(1), 1–39. <https://doi.org/10.46661/lexsocial.1165>

позбавляє судові органи адміністративного контролю над ними. Такі практики порушують фундаментальні принципи захисту даних, закріплені в Конвенції 108+¹⁹ та створюють серйозні ризики для інформаційної екосистеми правосуддя.

29. Універсальні комерційні сервіси генеративного штучного інтелекту, навіть якщо вони надаються на платній основі, не мають спеціалізованого навчання, необхідного для здійснення правового аналізу та підготовки судової документації. Такі інструменти, розроблені для загального споживчого використання, демонструють обмежене розуміння юридичної термінології, процесуальних вимог і тонких моделей аргументації, притаманних судовій діяльності. Відсутність інтеграції з перевіреними правовими базами даних підвищує ризик «галюцинацій», а саме генерування неіснуючих норм права або викривленого відтворення судової практики, що становить серйозну загрозу правовій визначеності.

30. У разі використання комерційних сервісів упровадження генеративного штучного інтелекту в системі правосуддя має здійснюватися виключно на підставі формалізованих договірних відносин між органами публічної влади та постачальниками сервісів генеративного штучного інтелекту. У разі використання внутрішніх (in-house) систем упровадження генеративного штучного інтелекту в системі правосуддя має здійснюватися на підставі внутрішнього нормативного документа щодо використання штучного інтелекту в системі правосуддя²⁰ або в межах комплексної системи інформаційної безпеки.²¹ Такі договори або внутрішні нормативні документи мають встановлювати комплексні рамки, спрямовані на врегулювання:

- запровадження суворих протоколів управління даними, що забезпечують запобігання несанкціонованій передачі даних третім особам;
- визначення обов'язкових процедур знищення даних після спливу встановлених строків їх зберігання;
- пряму заборону використання судових даних для повторного навчання моделей без прямо наданого дозволу;
- чітке розмежування відповідальності у разі порушень безпеки даних або збоїв у функціонуванні систем;

¹⁹ Рада Європи. (2018). Конвенція про захист осіб щодо автоматизованої обробки персональних даних [Модернізована Конвенція 108], Серія договорів Ради Європи – № 223. Страсбург.

²⁰ Наприклад: Державний технічний комітет з електронного адміністрування правосуддя Іспанії (State Technical Committee for the Electronic Judicial Administration). (2024). Внутрішній нормативний документ щодо використання штучного інтелекту в системі правосуддя. Неофіційний переклад англійською мовою доступний за посиланням: <https://www.administraciondejusticia.gob.es/cteaje/normativa-complementaria>

²¹ Наприклад: Королівський декрет Іспанії № 311/2022, яким регулюється Національна схема безпеки

- дотримання національного та європейського законодавства у сфері захисту даних;
- регулярне проведення аудитів безпеки та тестувань на проникнення;
- встановлення прозорих процедур реагування на інциденти;
- механізми постійного залучення заінтересованих сторін, зокрема шляхом проведення публічних консультацій, створення наглядових комітетів та здійснення періодичних оцінювань.

31. Спеціалізовані судові моделі генеративного штучного інтелекту є ключовою передумовою ефективного впровадження штучного інтелекту у сфері правосуддя. Правнича мова становить окрему технічну сферу зі спеціалізованою термінологією, яка часто істотно відрізняється від загальноновживаної. Такі терміни, як «consideration» (зустрічне задоволення в договірному праві), «prejudice» (упередженість або завдання шкоди), «public order» (публічний порядок) або «bona fides» (добросовісність), мають специфічне юридичне значення, що суттєво відрізняється від їх повсякденного розуміння. Така семантична складність зумовлює необхідність використання моделей генеративного штучного інтелекту, спеціально навчених на корпусах правничих текстів кожної окремої юрисдикції, з метою забезпечення коректного застосування фахової юридичної термінології та належного контекстуального тлумачення.

32. Індивідуально налаштовані судові системи генеративного штучного інтелекту можуть включати, зокрема:

- інтеграцію з офіційними правовими базами даних та перевіреними джерелами судової практики;
- навчання з урахуванням особливостей конкретної юрисдикції на правничих текстах, судових рішеннях і процесуальних документах;
- багатомовні можливості, що відображають мовне різноманіття європейських судових систем;
- спеціалізовані модулі для різних галузей права (цивільного, кримінального, адміністративного, конституційного);
- вбудовані системи перевірки посилань, які здійснюють звірку з офіційними правовими реєстрами;
- автоматизоване виявлення та маркування потенційних правових неузгодженостей або анахронізмів.

33. Суверенітет інфраструктури є ключовим чинником. Судові системи генеративного штучного інтелекту мають функціонувати на інфраструктурі, що перебуває під виключним контролем судових органів або інших органів

державної влади у разі використання відкритих моделей із відкритим вихідним кодом, розгорнутих на публічних серверах, або в межах співпраці з компаніями-постачальниками у разі використання комерційних моделей.

2.1.5. Забезпечення належного управління під час упровадження систем генеративного штучного інтелекту

34. Упровадження систем генеративного штучного інтелекту в судовому адмініструванні має здійснюватися на основі комплексного підходу, що охоплює поетапну стратегію впровадження з належними процесами управління змінами та пілотними етапами з подальшим оцінюванням і переглядом, зважений вибір моделей ліцензування (пропрієтарні технології або технології з відкритим вихідним кодом), безперервну підтримку з боку постачальників і вимоги щодо навчання, а також регулярне проведення аудитів, моніторинг і перегляд ефективності функціонування систем штучного інтелекту та пов'язаних із ними ризиків.

35. З метою забезпечення кращого розуміння технології, зниження потенційних бар'єрів для її впровадження та сприяння ефективному управлінню змінами рекомендується застосовувати поетапний, поступовий підхід до впровадження генеративного штучного інтелекту в системі правосуддя, диференційований залежно від складності та характеру завдань, що вирішуються:

Підготовчий етап: до початку будь-якого впровадження мають бути проведені оцінювання інституційної готовності, забезпечено базову підготовку персоналу, запроваджено протоколи безпеки та створено наглядові комітети.

Перший етап: обмежується базовими адміністративними та допоміжними функціями, які не мають безпосереднього впливу на ухвалення судових рішень. На цьому етапі основна увага має приділятися управлінню документами та обробці інформації, зокрема вилученню релевантних даних із неструктурованих документів, отриманню ключової інформації (імен, дат, сум) та їх автоматичному внесенню до заздалегідь визначених шаблонів. Також на цьому етапі може бути передбачене базове управління навантаженням шляхом аналізу історичних закономірностей і надання допомоги у розподілі справ. Крім того, можуть упроваджуватися базові інформаційні сервіси та інструменти допомоги користувачам, зокрема системи процесуальних роз'яснень зрозумілою мовою для користувачів і

спеціалізовані чат-боти для вирішення простих процесуальних запитів. Такі інструменти мають супроводжуватися застереженнями щодо пов'язаних ризиків. На цьому етапі також може розглядатися можливість несинхронного транскрибування судових засідань. Суди та органи, відповідальні за адміністрування правосуддя, мають оцінювати вплив на навантаження, здійснювати моніторинг рівня задоволеності користувачів і вимірювати час, необхідний для адаптації.

Другий етап: розширення до більш складних аналітичних функцій і функцій підтримки підготовки документів. Цей етап має охоплювати розширені можливості управління документами та обробки інформації, зокрема автоматичне формування протоколів судових засідань і встановлення зв'язків між аргументами в різних процесуальних документах. На цьому етапі також може бути передбачено підтримку підготовки процесуальних документів і судової документації, зокрема автоматичне створення стандартних процесуальних документів, допомогу в підготовці типових ухвал та перевірку узгодженості документів. Крім того, цей етап може включати створення сховищ судової практики, структурованих за видами справ і предметом спору, а також автоматизоване виявлення спірних питань шляхом порівняння позовних заяв і відзивів, у поєднанні з підготовкою тематичних і хронологічних покажчиків для об'ємних матеріалів справ. Аналіз судової практики та підтримка прийняття рішень мають запроваджуватися для попереднього аналізу апеляційних скарг з точки зору дотримання формальних і технічних вимог. Суди мають проводити A/B-тестування для порівняння ефективності, оцінювати потенційну упередженість рішень та аналізувати моделі використання систем.

Третій етап: запровадження більш просунутих аналітичних і консультативних функцій за умови збереження суворого судового нагляду. Реалізація цього етапу потребує позитивного оцінювання результатів попередніх етапів, надання окремого схвалення для кожного підетапу, а також наявності чітко визначених протоколів повернення до попереднього стану (rollback). Цей етап може бути додатково поділений на три підетапи:

Початковий підетап: зосереджується на дослідженні судової практики та впорядкуванні документів шляхом аналізу позовних заяв і відзивів із наданням пропозицій щодо релевантної судової практики, а також на виявленні й систематизованому упорядкуванні юриспруденції, що має значення для кожної конкретної справи та подається на розгляд суду. Допомога у

підготовці документів може включати формування структурованих стислих викладів судових рішень, процесуальних документів і апеляційних скарг за умови обов'язкової всебічної судової перевірки. Розширені функції управління справами можуть сприяти виявленню моделей аргументації у подібних позовах і визначенню проваджень, у яких можуть порушуватися аналогічні правові питання.

Проміжний підетап: передбачає розширене дослідження судової практики та впорядкування документів з метою допомоги у підготовці до здійснення правосуддя, шляхом надання систематизованих правових прецедентів і релевантної судової практики для розгляду судом. Допомога у підготовці документів може включати підготовку допоміжних матеріалів, зокрема звітів, структурних планів і ментальних карт, які подають узагальнений огляд позицій сторін, зокрема пунктів згоди та розбіжностей між ними.

Просунутий підетап: у межах високостандартизованих процедур із добре усталеною судовою практикою генеративний штучний інтелект може надавати допомогу шляхом здійснення комплексного аналізу судової практики з виявленням закономірностей у подібних справах, ідентифікації процесуальних прецедентів і потенційних незгодженостей для судового перегляду, а також шляхом підготовки проєктів документів із урахуванням релевантної судової практики. Крім того, він може допомагати у підготовці структурованих шаблонів для типових процесуальних документів на основі усталених судових практик, за умови обов'язкового здійснення аналізу людиною та ухвалення нею всіх правових і судових рішень, а також остаточного схвалення кожного результату.

36. Усі наведені вище підходи мають застосовуватися у формі пілотних проєктів, зокрема в процедурах судових наказів про стягнення, справах щодо вимог на визначені суми, усних проваджень у справах незначної вартості, справах про стандартизовані адміністративні штрафи, а також у трудових спорах щодо безспірних сум, у яких усі докази мають документальний характер.

37. Питання, що зачіпають основоположні права, стосуються неповнолітніх або недієздатних осіб чи передбачають застосування заходів, пов'язаних із

позбавленням свободи, потребують повного судового втручання з боку людини; у таких випадках використання генеративного штучного інтелекту має обмежуватися виключно технічними допоміжними функціями.

2.1.6. Усвідомлення того, як сторони використовують генеративний штучний інтелект

38. Суди та органи, що забезпечують здійснення правосуддя, повинні запровадити чіткі протоколи для моніторингу та регулювання використання генеративного штучного інтелекту правниками та сторонами процесу. Це може, зокрема, передбачати обов'язок сторін і їхніх представників повідомляти про використання систем генеративного штучного інтелекту під час підготовки процесуальних документів, проведення дослідження матеріалів справи або формування доказової бази – у випадках, коли це визнається доцільним.

39. Процесуальні правила повинні передбачати обов'язок повідомляти про використання генеративного штучного інтелекту під час підготовки справи, а також встановлювати належні санкції за неподання такого повідомлення або за надмірне покладання на згенерований штучним інтелектом контент без належної перевірки.

40. Суди повинні проявляти особливу пильність щодо різних форм неправомірної поведінки, пов'язаної з використанням штучного інтелекту, у тому числі: подання сфабрикованих доказів, неіснуючих правових посилань або згенерованого штучним інтелектом контенту, який не був належним чином перевірений кваліфікованими фахівцями; спроби маніпулювання системами штучного інтелекту шляхом використання прихованих запитів чи інструкцій, вбудованих у документи, зокрема невидимого тексту, нетипового форматування або мікроскопічних вбудованих інструкцій, спрямованих на компрометацію аналізу із застосуванням ШІ; а також використання генеративного штучного інтелекту для створення відео- або графічних доказів шахрайського характеру. Усі такі практики повинні бути врегульовані як на процесуальному, так і на кримінально-правовому рівні.

41. З метою забезпечення процесуальної справедливості та рівності сторін суди повинні гарантувати, щоб усі сторони мали рівний доступ до інформації про використання генеративного штучного інтелекту у провадженні та щоб жодна зі сторін не отримувала несправедливої переваги внаслідок нерозкритого використання допомоги штучного інтелекту.

2.1.7. Забезпечення обізнаності та навчання користувачів штучного інтелекту

42. Судді, прокурори, працівники апарату та весь судовий персонал повинні проходити належну підготовку з метою розуміння, управління та критичної оцінки інструментів генеративного штучного інтелекту, доступних їм у роботі та наявних на ринку. Спеціалізоване навчання та підвищення рівня обізнаності щодо потенціалу й обмежень таких інструментів сприятимуть етичному та ефективному використанню цієї технології за умови збереження вирішальної ролі людського судження у процесі прийняття рішень.

43. Запровадження штучного інтелекту в діяльність органів правосуддя повинно характеризуватися прозорістю та передбачуваністю. З метою забезпечення правової визначеності та послідовності судді не повинні бути зобов'язані здійснювати спеціальні дії для активації системи; її використання має бути заздалегідь визначене процесуальним законодавством.

44. Основні компетентності суддів повинні зосереджуватися на збереженні та застосуванні їхнього критичного судження — якості, притаманної судовій функції та очікуваної від неї, — з обов'язком ретельно оцінювати кожен результат, наданий штучним інтелектом. Це передбачає уникнення самозаспокоїливої або пасивної позиції щодо технологічних пропозицій і, натомість, формування критичного, а за потреби й незгодного ставлення до окремих альтернатив чи пропозицій, згенерованих системами генеративного штучного інтелекту.

45. У разі якщо судді будуть оснащені системами генеративного штучного інтелекту, з якими вони можуть безпосередньо взаємодіяти, вони повинні пройти спеціалізоване навчання щодо їх використання (зокрема з питань формування запитів (*prompt engineering*), контекстуалізації (*context engineering*) тощо). Судді мають набути ґрунтового технічного розуміння притаманних таким моделям обмежень, зокрема обсягу їх можливостей, сфер, у яких вони схильні давати збої, їхніх сильних сторін, а також типів запитів, які є технічно нежиттєздатними або недоречними.

46. Судді повинні мати здатність аналізувати та ставити під сумнів результати, сформовані генеративним штучним інтелектом, виявляючи можливі упередження, непослідовності або обмеження у логіці обґрунтування. Вони повинні розробляти критерії для оцінки того, коли і в який спосіб доцільно використовувати генеративний штучний інтелект у різних категоріях справ, а

також розвивати здатність ефективно переходити між традиційними джерелами (такими як бази даних судової практики та доктринальні статті) і системами генеративного штучного інтелекту з метою отримання найбільш релевантної та надійної інформації.

47. Суддів слід навчати уникати так званого «ефекту бульбашки», який виникає тоді, коли суддя, вражений уявною ефективністю та точністю систем генеративного штучного інтелекту, формує надмірну довіру до судової практики, звітів або узагальнень, які вони надають, що створює ризик відмови від самостійного звернення до оригінальних баз даних судової практики та ігнорування первинних документів.

48. Має бути розроблена політика безперервного навчання щодо нових інструментів генеративного штучного інтелекту в міру їх появи, а також щодо досягнень, яких сам штучний інтелект набуває в різних сферах, що впливають на здійснення правосуддя (зокрема генерація зображень, відео, голосів тощо). Судові установи повинні запровадити системні механізми для постійного інформування працівників судової системи про нові можливості штучного інтелекту та їхні потенційні наслідки для діяльності органів правосуддя. Це має охоплювати розуміння нових форм доказів, згенерованих штучним інтелектом, виявлення дипфейків і синтетичного контенту, а також адаптацію до методів цифрових маніпуляцій, що постійно розвиваються, які можуть впливати на судові провадження. Таке навчання повинно регулярно оновлюватися з урахуванням сучасного рівня розвитку технологій генеративного штучного інтелекту та їх застосування в правовому контексті. З метою забезпечення послідовності застосування можуть бути розглянуті стандартизовані навчальні програми.

2.2. Нормативні аспекти використання генеративного штучного інтелекту в судах

49. Запровадження систем генеративного штучного інтелекту в судовому середовищі потребує ґрунтовної оцінки їх відповідності основоположним правовим принципам, закріпленим у європейських правових системах та національних конституційних положеннях. З метою забезпечення дотримання цих правових принципів можуть застосовуватися інструмент CEPEJ для оцінки²² та методологія аналізу впливу HUDERIA²³.

²² Європейська комісія з питань ефективності правосуддя (CEPEJ). (4 грудня 2023 року). Інструмент для оцінки операціоналізації Європейської етичної хартії щодо використання штучного інтелекту в судових системах та їх середовищі. Рада Європи.

²³ Рада Європи, Комітет зі штучного інтелекту. (28 листопада 2024 року). Методологія оцінки ризиків та впливу систем штучного інтелекту з точки зору прав людини, демократії та верховенства права (методологія HUDERIA). Рада Європи.

2.2.1. Право на справедливий суд і суддівська незалежність (стаття ЕКПЛ)²⁴

6

2.2.1.1. Гарантії справедливого судового розгляду

50. Системи генеративного штучного інтелекту можуть негативно впливати на справедливість судового розгляду через непрозорі процеси ухвалення рішень, автоматизовані рекомендації, що не піддаються поясненню, або нерівні процесуальні переваги, які виникають унаслідок різного рівня доступу сторін до сучасних технологічних інструментів.²⁵

51. Технічні запобіжники, що впливають із цих положень, повинні включати впровадження пояснюваних систем штучного інтелекту,²⁶ забезпечення повної простежуваності процесів міркування,²⁷ рівний доступ усіх сторін до результатів, згенерованих штучним інтелектом, а також технічні механізми, які гарантують реальну рівність сторін.²⁸ Системи штучного інтелекту, що підтримують ухвалення судових рішень, повинні надавати результати, які є водночас обґрунтовано пояснюваними та надійними і зрозумілими для фахівців без технічної підготовки. Вони також повинні впроваджувати заходи для мінімізації неточностей, забезпечувати дотримання коректної правової логіки та забезпечення безпеки даних.

2.2.1.2. Принцип законності, правової визначеності та передбачуваності

52. Системи генеративного штучного інтелекту, які продукують так звані «правові галюцинації» або посиляються на неіснуючі правові норми, становлять серйозні ризики для правової визначеності та верховенства права.

53. Використання штучного інтелекту не повинно підривати передбачуваність і послідовність судових рішень. Джерела інформації, критерії та методології, що використовуються моделями генеративного штучного інтелекту, повинні бути ретельно задокументовані та чітко обґрунтовані з метою недопущення внесення елементів невизначеності або свавільності у

²⁴ Конвенція про захист прав людини і основоположних свобод від 4 листопада 1950 року, ETS № 5, 213 U.N.T.S. 221. https://www.echr.coe.int/documents/d/echr/convention_ENG

²⁵ Пункт 1 статті 6 ЕКПЛ гарантує право на «(...) справедливий і публічний розгляд справи упродовж розумного строку незалежним і безстороннім судом, встановленим законом»

²⁶ Рівень розуміння того, яким чином система на основі штучного інтелекту дійшла до певного результату; див. ISO/IEC TR 29119-11:2020.

²⁷ Вимога до організацій забезпечувати ведення детальних журналів роботи систем штучного інтелекту, зокрема щодо вхідних даних, рішень моделей, помилок і взаємодії з користувачами; див. ISO/IEC 42001:2023.

²⁸ Інструмент СЕРЕУ для оцінки ідентифікує як ризик № 8: «Ризик отримання несправедливої переваги однією зі сторін судового розгляду».

судове провадження.

54. Технічні запобіжники повинні включати автоматизовану перевірку правових посилань, виключне використання сертифікованих офіційних правових джерел, а також системи сповіщення щодо згенерованого контенту, який не може бути перевірений за авторитетними правовими базами даних.²⁹

2.2.1.3. Виключність судової влади

55. Юрисдикційні повноваження належать виключно суддям відповідно до конституційних і законодавчих положень. Прийняття рішень, правове обґрунтування та оцінка доказів не повинні делегуватися системам генеративного штучного інтелекту; такі системи можуть застосовуватися лише для допоміжних або підготовчих завдань, які підтримують, але ніколи не замінюють людське судження судового органу.

2.2.1.4. Обов'язкова відповідальність людини та нагляд

56. Судді зберігають остаточну та виключну відповідальність за всі судові рішення незалежно від будь-якої технологічної допомоги, що застосовується у процесі їх ухвалення. Вони повинні повною мірою розуміти обмеження будь-якої системи генеративного штучного інтелекту, яка використовується, та ретельно перевіряти всі результати, згенеровані таким штучним інтелектом, перед їх включенням до судових актів або рішень.

2.2.1.5. Суддівська незалежність і необов'язковий характер використання штучного інтелекту

57. Суддівська незалежність³⁰ може бути підірвана у випадках, коли судді зазнають прямого або опосередкованого тиску щодо дотримання рекомендацій, згенерованих штучним інтелектом, що обмежує їхню свободу судження, тлумачення права та автономність процесу ухвалення рішень. **Вона** також передбачає свободу від технологічного примусу або системних упереджень, запроваджених через рекомендації штучного інтелекту. Судді повинні зберігати право відмовитися від отримання пропозицій, сформованих штучним інтелектом.

58. Рекомендації, пропозиції або результати, згенеровані системами штучного інтелекту, ніколи не є обов'язковими для суддів, так само як і їх

²⁹ Інструмент СЕРЕУ для оцінки ідентифікує як ризик № 11: «Ризик генерування та використання неіснуючих правових норм генеративним штучним інтелектом».

³⁰ Пункт 1 статті 6 ЄКПЛ вимагає розгляду справи «незалежним і безстороннім судом, встановленим законом».

використання не може бути обов'язковим, що забезпечує збереження суддівської незалежності у тлумаченні права, обґрунтуванні та ухваленні рішень. Судді не повинні бути зобов'язані спеціально обґрунтовувати відхилення від пропозицій чи рекомендацій, згенерованих системами штучного інтелекту.

59. Проєктування моделей генеративного штучного інтелекту не повинно надавати пріоритету максимізації їх переконливості з метою запобігання ризику неправомірного психологічного впливу та упередженості автоматизації, які можуть негативно позначатися на суддівській неупередженості та свободі під час оцінки висновків, згенерованих штучним інтелектом.³¹

60. Технічні запобіжники повинні бути пропорційними ризикам, пов'язаним із відповідними системами, та включати проєктування моделей генеративного штучного інтелекту виключно як необов'язкових допоміжних інструментів, усунення будь-яких механізмів, що зобов'язують або чинять тиск на суддів з метою обґрунтування відхилення від алгоритмічних пропозицій, проведення ретельної оцінки до будь-якого впровадження інструментів штучного інтелекту, здійснення незалежних технічних аудитів, а також запровадження механізмів безперервного навчання, спрямованих на запобігання упередженості автоматизації та збереження автономного людського судження.³²

2.2.1.6. Прозорість та належне обґрунтування

61. Використання генеративного штучного інтелекту в будь-якому аспекті ухвалення судових рішень повинно бути чітко задокументоване.³³ Сторони мають право знати, які саме елементи судового рішення були підготовлені із застосуванням систем генеративного штучного інтелекту.

2.2.1.7. Право на доступ до суду

62. Це основоположне право може зазнавати загрози внаслідок

³¹ Салві, Ф., Орта Рібейру, М., Ґаллотті, Р., та Вест, Р. (2024). Про розмовну переконливість великих мовних моделей: рандомізоване контрольоване дослідження. Препринт arXiv, arXiv:2403.14380.

³² Інструмент CEPEJ для оцінки ідентифікує як ризик № 12: «Ризик позбавлення судді повноважень та обмеження його відповідальності внаслідок використання непояснюваного штучного інтелекту». Водночас рамкова методологія HUDERIA забезпечує можливість періодичної оцінки потенційних негативних впливів на суддівську автономію.

³³ Європейська комісія з питань ефективності правосуддя. (31 січня 2025 року). Міркування Консультативного органу з питань штучного інтелекту (AIAB) щодо використання штучного інтелекту в судових системах (CEPEJ-AIAB(2025)1Rev5).

використання автоматизованих систем, які замінюють або ускладнюють доступ до ухвалення судових рішень людиною, що впливає на право на розгляд і вирішення судових справ людиною-суддею.

63. Технічні запобіжники повинні гарантувати реальний і змістовний доступ до судді-людини, обмежувати використання генеративного штучного інтелекту виключно допоміжними функціями та забороняти повну автоматизацію ухвалення судових рішень.³⁴

2.2.2. Право на свободу та особисту недоторканність (стаття 5 ЄКПЛ)

64. Рішення, що впливають на особисту свободу, потребують найвищого рівня людського нагляду та належного обґрунтування. Системи генеративного штучного інтелекту, які впливають на рішення щодо затримання, тримання під вартою до суду або застосування запобіжних заходів без належної перевірки людиною та прозорого обґрунтування, становлять істотні ризики для основоположних свобод.³⁵

65. Запровадження штучного інтелекту у здійснення судових функцій має відбуватися за поетапним, поступовим підходом і на початковому етапі обмежуватися адміністративними завданнями або стандартизованими процедурами з передбачуваними результатами. Точність, ефективність і безпека систем генеративного штучного інтелекту повинні підлягати ретельній оцінці перед розширенням їх застосування на більш чутливі судові функції.

66. Технічні запобіжники повинні включати використання систем генеративного штучного інтелекту, спрямованих на заохочення альтернатив позбавленню волі, обов'язкове надання пояснень у зрозумілій для людини формі щодо рекомендацій, згенерованих штучним інтелектом і таких, що впливають на питання свободи, а також всебічний людський перегляд кожного рішення, ухваленого із застосуванням штучного інтелекту, у випадках, коли може бути застосоване позбавлення волі.

2.2.3. Право на повагу до приватного життя та захист персональних даних (стаття 8 ЄКПЛ)

67. У разі використання систем генеративного штучного інтелекту має бути

³⁴ Інструмент CEREJ для оцінки ідентифікує як ризик № 6: «Ризик заміщення доступу до судді системами штучного інтелекту».

³⁵ Пункт 1 статті 5 ЄКПЛ гарантує право на свободу за винятком лише чітко визначених законом випадків. Пункт 2 статті 5 встановлює право бути поінформованим «про підстави свого затримання». Пункт 4 статті 5 гарантує право на судовий перегляд законності тримання під вартою.

забезпечено дотримання права на приватність та захист персональних даних. Усі технології та дані, що застосовуються, повинні перебувати під виключним контролем судової влади з метою запобігання несанкціонованому доступу, витокам або зовнішньому втручання, а також збереження інституційного суверенітету над інформацією. Судові дані повинні залишатися під контролем судових органів. Суди мають зберігати спроможність змінювати постачальників систем штучного інтелекту без втрати функціональності, а також право перевіряти алгоритми та навчальні дані шляхом постійних аудиторських процедур.

68. Технічні запобіжники можуть включати обов'язкову локальну псевдонімізацію персональних даних до їх завантаження в систему штучного інтелекту, а також виконання моделей штучного інтелекту на інфраструктурі, що перебуває під контролем державних органів. У разі використання комерційно доступної інфраструктури це має здійснюватися за посередництвом формалізованих договірних домовленостей між органами публічної влади та постачальниками технологічних послуг, із запровадженням суворих протоколів захисту даних.³⁶

2.2.4. Право на ефективний засіб правового захисту (стаття 13 ЄКПЛ)

69. Автоматизовані системи на основі штучного інтелекту не повинні перешкоджати ефективному доступу до засобів правового захисту або судового розгляду.³⁷

70. Технічні запобіжники повинні забезпечувати обов'язковий перегляд рішень людиною у випадках використання штучного інтелекту, а також передбачати спеціальні процедури правового захисту для оскарження автоматизованих рішень або рішень, на формування яких вплинуло використання штучного інтелекту.

2.2.5. Заборона дискримінації (стаття 14 ЄКПЛ та Протокол № 12)

2.2.5.1. Принцип недискримінації

71. Алгоритмічні упередження в системах генеративного штучного

³⁶ Інструмент СЕРЕУ для оцінки ідентифікує як ризик № 2: «Ризик розкриття персональних даних або комерційної таємниці». Рамкова методологія аналізу впливу на приватність HUDERIA безпосередньо охоплює ризики несанкціонованого розкриття чутливих даних, зокрема через методологію аналізу ризиків COBRA.

³⁷ Стаття 13 ЄКПЛ встановлює, що «кожен, чий права і свободи, викладені в цій Конвенції, було порушено, має право на ефективний засіб правового захисту перед національним органом, навіть якщо порушення було вчинене особами, які діяли в офіційній якості».

інтелекту становлять істотні ризики відтворення або посилення наявних форм дискримінації за захищеними ознаками, що потенційно може призводити до порушення основоположних принципів рівності.³⁸

72. Технічні запобіжники повинні включати періодичні аудити упередженості, очищення та збалансування даних, застосування під час навчання моделей регуляризаційних механізмів забезпечення справедливості (зокрема обмежень на кшталт демографічної паритетності або рівності шансів), проведення контрфактичного та A/B-тестування, а також безперервний моніторинг показників дискримінації.³⁹

2.2.5.2. Нейтральність та активне запобігання упередженості

73. Системи штучного інтелекту повинні періодично оцінюватися з метою виявлення та пом'якшення упередженості, щоб гарантувати, що вони не відтворюють наявні форми дискримінації та не створюють нові прояви нерівності у процесі ухвалення судових рішень. Обов'язковим є систематичний моніторинг за всіма захищеними ознаками.

2.2.6. Додаткові запобіжники

2.2.6.1. Пояснюваність та прозоре обґрунтування

74. Системи штучного інтелекту, що використовуються у судовій сфері, повинні надавати чіткі й зрозумілі пояснення шляху, яким було отримано кожен результат або рекомендацію, щоб суб'єкти здійснення правосуддя могли критично їх перевіряти. З цією метою такі системи повинні розкривати свої внутрішні процеси виведення у прозорий та придатний до аудиту спосіб, забезпечуючи змістовну відповідність і уникаючи формальної імітації дотримання вимог, що маскує їхню реальну логіку.⁴⁰ Системи мають бути спроможні вказувати конкретні джерела кожного згенерованого твердження. Слід використовувати лише інтерпретовані моделі, які розкривають свої

³⁸ Стаття 14 ЄКПЛ встановлює, що «користування правами і свободами, викладеними в цій Конвенції, повинно бути забезпечене без дискримінації за будь-якою ознакою, такою як стать, раса, колір шкіри, мова, релігія, політичні чи інші переконання, національне або соціальне походження, належність до національної меншини, майновий стан, народження або інший статус».

³⁹ Інструмент CEPEJ для оцінки ідентифікує як ризик № 10: «Ризик дискримінації або посилення дискримінації». Процес залучення заінтересованих сторін у межах методології HUDERIA дає змогу виявляти диференційований вплив на вразливі групи.

⁴⁰ Грінблатт, Р., Денісон, К., Райт, Б., Роджер, Ф., Мак-Діармід, М., Маркс, С., Тройтляйн, Дж., Белонакс, Т., Чен, Дж., Дювенот, Д., Хан, А., Майкл, Дж., Міндерманн, С., Перес, Е., Петріні, Л., Уесато, Дж., Каплан, Дж., Шлегеріс, Б., Боумен, С. Р., та Г'юбінґер, Е. (2024). Імітація узгодженості у великих мовних моделях [препринт arXiv:2412.14093v2]. <https://doi.org/10.48550/arXiv.2412.14093>

внутрішні процеси ухвалення рішень.

2.2.6.2. Пропорційність і обережність у впровадженні генеративного штучного інтелекту

75. Запровадження штучного інтелекту у здійснення суддівських функцій має відбуватися за поетапним, поступовим підходом і на початковому етапі обмежуватися адміністративними завданнями або стандартизованими процедурами з передбачуваними результатами. Точність, ефективність і безпека систем генеративного штучного інтелекту повинні підлягати ретельній оцінці перед розширенням їх застосування на більш чутливі суддівські функції.

2.2.6.3. Відповідальність держави за шкоду, завдану використанням генеративного штучного інтелекту

76. Держава повинна нести відповідальність за будь-яку шкоду, заподіяну внаслідок помилок судді або збоїв у роботі систем генеративного штучного інтелекту, відповідно до чинних принципів державної відповідальності. Ані автономність, ані технологічна складність систем генеративного штучного інтелекту не можуть слугувати підставою для звільнення держави від обов'язку забезпечити відшкодування шкоди.

2.2.6.4. Принцип процесуальної справедливості та ефективності

77. Запровадження штучного інтелекту у здійснення правосуддя не повинно призводити до послаблення процесуальних гарантій або зниження рівня ефективного судового захисту; навпаки, воно має посилювати процесуальні права та ефективність правосуддя. Зокрема, воно повинно поліпшувати доступ до суду, забезпечувати кращий доступ до доказів або вдосконалювати засоби належного представлення позиції у справі. Крім того, впровадження генеративного штучного інтелекту має призводити до вимірюваних поліпшень щодо строків реагування, чіткості комунікації, послідовності судових рішень та ефективності управління справами. Суди повинні зберігати спроможність повернутися до попередніх систем у разі, якщо генеративний штучний інтелект об'єктивно не поліпшує якість судових послуг. Регулярні оцінювання мають визначати дотримання цього принципу.⁴¹

⁴¹ Показники оцінювання повинні включати опитування щодо задоволеності користувачів, порівняльний аналіз показників ефективності із системами, що застосовувалися до впровадження штучного інтелекту, зворотний зв'язок від заінтересованих сторін, зокрема суддів, працівників судової системи, адвокатів і громадян, а також оцінку поліпшень у дотриманні процесуальних строків.

3. ГЛОСАРІЙ

Принцип технологічної нейтральності

Загалом розуміється як свобода фізичних осіб і організацій обирати найбільш доцільні та відповідні їхнім потребам і вимогам технології для розроблення, придбання, використання або комерціалізації без залежності від конкретних знань, а також від інформації чи даних, що з ними пов'язані.⁴²

Регуляризатори справедливості у процесі модельного навчання та налаштування

Специфічні математичні методи, зокрема:

- Регуляризатори демографічної паритетності – методи, що обмежують відмінності у частоті прогнозованих результатів між різними демографічними групами.
- Регуляризатори рівності шансів (equalized odds) – підходи, спрямовані на наближення показників істинно позитивних і хибнопозитивних результатів між захищеними групами.
- Обмеження індивідуальної справедливості – методи, що забезпечують подібне ставлення до осіб з подібними характеристиками.
- Адверсаріальне усунення упередженості (adversarial debiasing) – використання змагальних моделей для зменшення або усунення інформації про захищені ознаки з внутрішніх представлень моделі.

⁴² Див.: https://icannwiki.org/Technology_neutrality