

Іван Городиський

LEX EX MACHINA

**Правові, етичні та світоглядні
виклики застосування штучного
інтелекту в правосудді**

Квітень 2023

LEX EX MACHINA: Правові, етичні та світоглядні виклики застосування штучного інтелекту в правосудді. Аналітичний документ / І.М. Городиський. Львів: Центр Дністрянського, Український католицький університет 2023. 21с.

В аналітичному документі здійснено огляд ключових правових, етичних та світоглядних проблем застосування інструментів штучного інтелекту в судочинстві. Метою є привернення уваги до цієї проблеми в контексті тематики цілісного людського розвитку та започаткування ширшої експертної дискусії, зокрема в Україні, яка уже невдовзі може стикнутися з викликами, пов'язаними з впровадженням таких інструментів.

Автор: Іван Городиський, к.ю.н., адвокат, директор-засновник Школи права УКУ (2014-2020), директор Центру Дністрянського, член Правління Асоціації правників України.

Публікацію підготовлено в рамках цільової науково-дослідної програми Цілісного людського розвитку (ЦЛР) Українського католицького університету.

© Український католицький університет

© І.М. Городиський

Зміст

Виконавче резюме	4
Опис проблеми	5
Досвід застосування ШІ в правосудді.....	8
Аналіз наявних рішень	10
Світоглядні виклики застосування ШІ в правосудді.....	13
Можливі рішення	18
Висновки і рекомендації	20

Виконавче резюме

Застосування штучного інтелекту в правовій сфері, зокрема у правосудді, є не найпомітнішим, але дуже важливим аспектом використання цих технологій. Нижче буде здійснено огляд ключових правових, етичних та світоглядних проблем застосування інструментів штучного інтелекту в судочинстві. Метою є привернення уваги до цієї проблеми в контексті тематики цілісного людського розвитку та започаткування ширшої експертної дискусії, зокрема в Україні, яка уже невдовзі може стикнутися з викликами, пов'язаними з впровадженням таких інструментів.

У першій частині буде розглянуто проблему використання технологій штучного інтелекту в правосудді в цілому, її витоки та виклики поточного моменту. *У другій частині* буде проаналізовано підходи до визначення юридичних, моральних та етичних принципів врегулювання цього процесу, з особливою увагою до напрацювань Католицької Церкви, яка бере активну участь у суспільному дискурсі на тему розвитку штучного інтелекту. *У третій частині* на основі проведеного дослідження буде окреслено можливі шляхи розв'язання проблем, пов'язаних із застосуванням таких технологій у судочинстві. У Висновках та рекомендаціях буде підсумовано проведене дослідження та сформульовано конкретні пропозиції з цього приводу.

При підготовці документа були проаналізовані як офіційні проекти врегулювання згаданих проблем, так і теоретичні підходи, що дають їм оцінку. Документ адресований експертним середовищам, які займаються використанням технологій у сфері права з юридичної точки зору, а також тим, кого цікавить їхній вплив на суспільство в цілому та на цілісний людський розвиток.

Опис проблеми

У наш час, у багатьох суспільствах, в тому числі українському, авторитет професії юриста в будь-якій сфері: приватній практиці, судочинстві чи правоохоронній діяльності – неабияк підірваний. Така ситуація часто має підстави в нездатності самих юристів забезпечити реальне верховенство права і захист особистих прав і свобод. Одним із наслідків цього є дискусії про впровадження технологій в юридичну діяльність і, в перспективі, заміну юристів роботами/штучним інтелектом. При цьому часто ідеться не просто про автоматизацію процесів, а й про те, щоб передати технологіям найважливіші юридичні функції, зокрема правосуддя.

Штучний інтелект (ШІ), а точніше технології, які ґрунтуються на його ідеї та принципах, є однією із найпопулярніших зараз тем. Поява і масове використання сервісів Chat GPT/GPT-4 дали новий поштовх дискусіям про суспільний розвиток у контексті розвитку технологій ШІ.

Навколо ШІ існує багато міфів, його сучасне сприйняття сформоване значною мірою мистецтвом, зокрема літературою та кінематографом, де він виступає як система, що інтелектом у багато разів перевищує людину. І нині саме в таких сюжетах вбачають витoki ідей для вироблення засад етичного чи правового врегулювання розробки та застосування технологій ШІ.

Ідея штучного інтелекту, в принципі, дуже нагадує біблійну історію творіння. Людство завжди прагнуло пізнати таїнство творіння життя і самому стати творцем. Міф про Пігмаліона і Галатею, легенда про Голема чи роман Мері Шеллі «Франкенштейн» – це лише перші спроби людства припустити, як може виглядати штучне творіння життя людиною.

Однак у згаданих історіях мова йшла передовсім про творення тілесного життя, де розум був присутній «за замовчуванням», практично відсутній або вторинний. Науковий розвиток дав змогу переосмислити ідею творіння: після Другої світової війни, ще на ранніх стадіях розвитку комп'ютерних технологій, на ідею творіння зуміли подивитися під іншим кутом: творіння не просто тіла, а саме розуму. І якщо Бог створив людину за своїм «образом і подобою», то людина хоче творити розум не лише за подобою людського, а й кращий за нього.

Формальний юридичний дискурс, а відповідно й понятійний апарат щодо штучного інтелекту тільки зароджується. Наприклад, Європейська комісія з питань ефективності правосуддя Ради Європи (European Commission for the Efficiency of Justice, CEPEJ), визначає ШІ як «набір наукових методів, теорій і прийомів, метою яких є відтворення за допомогою машини когнітивних здібностей людини»¹. Це визначення багато в чому повторює те, яке дали ШІ

1 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. Council of Europe; European Commission for the Efficiency of Justice (CEPEJ): Strasbourg 2018. P. 69.

автори цього терміну на чолі з Джоном МакКарті ще у 1955 році: «здатність машин поводитися таким чином, що його можна було б назвати розумним, якби людина поводитися таким чином»².

Дослідники ШІ виділяють два його види: слабкий штучний інтелект (weak AI, narrow AI) – автоматизовані технології, які здатні приймати конкретні рішення з вузькоспеціалізованих питань, та сильний штучний інтелект (strong AI, artificial general intelligence, AGI) – технології, створені за моделлю свідомості людини, здатні до саморозвитку, когнітивні здібності яких можна порівняти з людськими.

Такі сервіси як Bard, Bing AI Chat, ChatGPT/GPT4, DALL-E, Midjourney часом виділяють в окрему категорію генеративного (мовленнєвого) ШІ (generative AI), яку часом класифікують як вид сильного ШІ. Однак попри значний розвиток когнітивних функцій, навряд чи можна провести знак рівності між цими прикладами і людською свідомістю.

Тепер і, ймовірно, у найближчому майбутньому ми матимемо справу зі слабким штучним інтелектом. У цьому документі ми будемо використовувати термін «технології ШІ», оскільки йтиметься переважно про слабкий штучний інтелект, із зарахуванням Chat GPT та його аналогів саме до цієї категорії.

Ті технології ШІ в правосудді, які уже застосовуються, – т.зв. технології прогнозованого або предиктивного правосуддя (англ. predictive justice) – підпадають під визначення саме слабого ШІ. Вони передбачають створення алгоритмів, що «вчаться» на рішеннях, прийнятих судьями-людьми.

Вже є прецеденти застосування технологій прогнозованого правосуддя як у сфері офіційного судочинства, так і на рівні сервісів приватних компаній, які допомагають спрогнозувати рішення суду. В приватній сфері активно розвиваються стартапи із застосування технологій ШІ для прогнозування рішень в судових справах, наприклад Lex Machina чи Ravel Law³.

Дослідники застосування ШІ в правосудді виділяють цілу низку перспектив, які ця практика може мати. Марія Димітрук називає, зокрема, такі: прискорення судового розгляду; уніфікація судової практики; підвищення економічної ефективності судової системи⁴. Голландська суддя Дорі Рейлінг говорить про використання ШІ в судочинстві з метою обробки інформації (organising information); консультування за запитом (advise); прогнозування (predictions)⁵.

2 McCarthy J., Minsky M.L., Rochester N., Shannon C.E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. August 31, 1955. URL : <https://www.jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> (date requested : 12.04.2023)

3 AI driven lawtech startups changing the legal landscape. URL : <https://outsideinsight.com/insights/4-ai-driven-lawtech-startups-changing-the-legal-landscape/> (date requested : 12.01.2022)

4 Dymitruk M. Ethical artificial intelligence in judiciary. 22nd International Legal Informatics Symposium IRIS. Salzburg, Austria 2019. Pp. 9.

5 Reiling A. D. (Dory), Courts and Artificial Intelligence. International Journal for Court Administration. 2020. Vol. 11, Iss. 2. URL : <https://ssrn.com/abstract=3736411> (date requested : 12.04.2023)

Перспективи використання технологій ШІ для здійснення правосуддя чи принаймні в ході його здійсненні зараз активно досліджуються. Наприклад, у 2020 році ЮНЕСКО провело опитування професійних суб'єктів судового процесу щодо ШІ та верховенства права (Survey of Judicial Operators on AI and the Rule of Law). Його учасниками стали 1265 осіб із понад 100 країн, 35% з яких були суддями. 85% опитаних висловили зацікавленість у використанні інструментів штучного інтелекту в правових системах, у тому числі у здійсненні правосуддя, проведенні розслідувань та в діяльності правоохоронних органів⁶.

Наразі складно спрогнозувати перспективи змін у здійсненні правосуддя саме застосуванням технологій прогнозованого правосуддя. На думку італійського дослідника та одного із розробників Європейської етичної хартії щодо використання штучного інтелекту у судочинстві Франческо Контіні, поки що «незрозуміло, чи ця тенденція веде до кращих рішень, чи підриває належне функціонування системи [правосуддя]»⁷.

Він наголошує на таких моментах, як підзвітність та здійснення контролю за алгоритмами. У разі застосування ШІ в правосудді, стверджує Контіні, «принцип обережності слід мати на увазі до тих пір, поки це питання не буде вирішене з технічної та інституційної точки зору»⁸.

Однак уже зараз робиться оцінка ризиків та перші кроки для їх попередження. Для цього визначаються правові стандарти розробки та застосування технологій ШІ в правосудді.

6 JJudicial operators in 100 countries express interest in learning about AI and the rule of law with UNESCO. URL : <https://en.unesco.org/news/judicial-operators-100-countries-express-interest-learning-about-ai-and-rule-law-unesco> (date requested : 12.01.2022)

7 Contini F. Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries? URL : https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html (date requested : 12.01.2022)

8 Contini F. Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries? URL : https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html (date requested : 12.01.2022)

Досвід застосування ШІ в правосудді

Хоча поява ChatGPT/GPT4 викликала дискусії щодо саме їх застосування у сфері юридичної практики і правосуддя, вони не стали першими технологіями ШІ, які застосовуються у цій сфері. У світі вже є прецеденти використання систем «з елементами ШІ» для відносно вузької сфери питань в галузі судочинства - розрахунку ймовірності рецидиву в осіб, які скоїли злочини, та оцінка результатів такого розрахунку.

Наприклад, у Британії з 2017 р. поліція у тестовому режимі використовує систему HART (англ. Harm Assessment Risk Tool), розроблену в Кембриджському університеті. У США є COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) – система, розроблена приватною компанією, яка вже застосовується у штатах Вісконсин, Нью-Йорк, Каліфорнія та ін. Результати їх застосування не є провальними: наприклад, HART при оцінці ймовірності рецидиву продемонстрував точність близько 90%. Проте в обох випадках звучала й критика, зокрема через недоступність інформації про код і алгоритми програм. Викликає застереження і те, що результати демонструють тенденцію до дискримінації – за соціальною ознакою в Британії або ж за расовою у США.

Американська неурядова організація ProPublica у своєму звіті про практику застосування COMPAS стверджує, що алгоритм цієї системи «позначав афроамериканських підсудних як майбутніх злочинців, помилково визначаючи їх такими майже вдвічі частіше, ніж білих підсудних». Ілюструючи це мовою цифр – афроамериканські підсудні, щодо яких було визначено високий ризик рецидиву, у 44,9% не вчиняли його, тоді як білі обвинувачені, щодо яких було визначено низький ризик рецидиву, вчиняли його у 47,7% випадків⁹. Фактично, ці приклади підтверджують тезу, яка вже раніше висловлювалася фахівцями з цієї проблематики, що: «нейтральність алгоритмів є міфом, оскільки їх творці свідомо чи ненавмисно переносять у них власні системи цінностей»¹⁰.

Характерно, що найактивніше дискусії щодо застосування ШІ в правосудді ведуться у двох типах суспільств: перехідних і авторитарних. Перші – це суспільства, які відчують особливі проблеми із доброчесністю та професіоналізмом у судовій системі та вбачають у ШІ шанс для виправлення ситуації. Наприклад, у 2020 році в Україні Міністр юстиції Денис Малюська виступив із ініціативою запуску т.зв. «Кассандри» – програмного забезпечення «з елементами штучного інтелекту» для підготовки характеристики обвинувачених, яка має враховуватися при винесенні вироку¹¹.

9 Machine Bias. URL : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (date requested : 12.01.2022)

10 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. Council of Europe; European Commission for the Efficiency of Justice (CEPEJ): Strasbourg 2018. Pp. 57.

11 Мін'юст запускає штучний інтелект для аналізу злочинців. URL : <https://suspilne.media/62673-minust-stvoriv-stucnij-intelekt-dla-analizu-zlocinciv/> (дата звернення : 12.01.2022)

Однак не лише перехідні суспільства активно розглядають можливість впровадження ШІ в правосудді: піонерами в цьому є авторитарні режими, які розглядають ці технології як інструмент контролю у сфері правосуддя. Наприклад, цілу низку можливостей зараз вивчає влада Китаю. Деякі з них, як наприклад впровадження т.зв. supervision platform в суді в місті Ібінь, викликають застереження з боку суддів, оскільки вони побоюються, що це стане інструментом контролю за ними¹².

Приклади HART, COMPAS та ін. здатні не тільки проілюструвати одну з проблем, пов'язаних із застосуванням технологій ШІ в правосудді на цьому етапі. Метою є вказати на інший імовірний ризик: формою усунення цієї проблеми може бути штучне «збалансування» алгоритмів на користь однієї суспільної групи або ж для «позитивної дискримінації» іншої. Тобто фактично втратиться сенс та ідея ШІ як автономної системи, а негативні явища, що існують у правосудді, будуть викривлені в інший спосіб і далі існуватимуть. Відкритим є також питання, чи зможуть рішення, прийняті такими відкоригованими алгоритмами, вважатися справедливими і правосудними.

Та все ж дискусії про застосування технологій ШІ в правосудді поки що сфокусовані на обговоренні самої можливості їх використання у цій сфері, а також перспектив, які воно може відкрити, і на проектуванні правового регулювання. Однак моральні та етичні засади цього процесу наразі активно не обговорюються.

12 Automation and Digitalization of Justice in China's Smart Court Systems. URL : <https://jamestown.org/program/automation-and-digitalization-of-justice-in-chinas-smart-court-systems/> (date requested : 12.01.2022)

Аналіз наявних рішень

Тема ШІ та його впливу на майбутнє людства викликає занепокоєння практично від самого початку появи цього терміну та висвітлення його в медіа та мистецтві. Такі популярні приклади як «Матриця», «Термінатор» чи «Світ Дикого Заходу» та низка інших, окрім цікавих сюжетів, для багатьох є ще й застереженням про небезпеку відсутності контролю за цими технологіями, якими б далекими вони не видавалися.

Заклики до обачності в питаннях розробки, впровадження та застосування ШІ висловлюються і провідними науковими і технологічними лідерами. Наприклад, Стівен Гокінг у 2014 році в інтерв'ю BBC похмуро передбачав, що «Розвиток сильного штучного інтелекту може означати кінець людської раси. Люди, обмежені повільною біологічною еволюцією, не змогли б конкурувати й були б витіснені»¹³.

Своєю чергою, Ілон Маск ще у 2017 році називав ШІ «найбільшим ризиком, з яким ми стикаємося як цивілізація», і закликав бути проактивними у розробці його правового регулювання¹⁴. А вже у 2023 році, на хвилі уваги до ChatGPT/GPT-4, Маск, Стівен Возняк та низка інших технологічних підприємців та мислителів закликали мінімум на 6 місяців призупинити розробку систем ШІ, що перевершують за потужністю GPT-4¹⁵.

Ці заклики можуть видаватися перебільшеними або ж такими, за якими стоять приховані мотиви, зокрема комерційні. Однак вони свідчать, що в перспективі розробка, поширення, впровадження та застосування ШІ у всіх сферах будуть врегульовані в принципі, і у сфері правосуддя – зокрема.

Дискусії про формальне врегулювання використання технологій ШІ в правосудді та юридичній практиці ведуться приблизно з середини 2010-х років. На важливості цього питання наголошують і професійні спільноти. Наприклад, Американська асоціація юристів (American Bar Association) у 2019 році закликала суди та правників вирішити етичні та правові проблеми, пов'язані з використанням ШІ у сфері права, зокрема щодо:

(1) упередженості, можливості обґрунтування та прозорості автоматизованих рішень, прийнятих ШІ;

(2) етичного та корисного використання ШІ;

13 Stephen Hawking warns artificial intelligence could end mankind. BBC. December, 02, 2014. URL : <https://www.bbc.com/news/technology-30290540> (date requested : 13.04.2023)

14 Condliffe J. Elon Musk Urges U.S. Governors to Regulate AI Before "It's Too Late". MIT Technology Review. July 17, 2017. URL : <https://www.technologyreview.com/2017/07/17/150434/elon-musk-urges-us-governors-to-regulate-ai-before-its-too-late/> (date requested : 13.04.2023)

15 Nolan B. Elon Musk and more than 1,000 people sign an open letter calling for a pause on training AI systems more powerful than GPT-4. Insider. URL : <https://www.businessinsider.com/elon-musk-ai-pause-open-ai-gpt4-powerful-development-2023-3> (date requested : 13.04.2023)

(3) засобів контролю та нагляду за ШІ та його постачальниками¹⁶.

Окремі заходи вживаються на рівні національних законодавств. Наприклад, Франція у 2019 році на офіційному рівні заборонила використання технологій «прогнозованого правосуддя» в судочинстві¹⁷. До цього спонукала, зокрема, позиція Конституційної ради Франції, яка висловила занепокоєння, що використання такої аналітики рішень може полегшити стратегію вибору судів і суддів, що, ймовірно, змінить саму природу функціонування правосуддя¹⁸.

В Україні наразі відсутні норми щодо регулювання застосування ШІ в правосудді. Проте в затвердженій Кабінетом міністрів Концепції розвитку штучного інтелекту в Україні є розділ «Правосуддя», який, зокрема, допускає «винесення судових рішень у справах незначної складності (за взаємною згодою сторін) на основі результатів аналізу, здійсненого з використанням технологій штучного інтелекту, стану дотримання законодавства та судової практики». Однак ані ця Концепція, ані План заходів до неї не містять конкретних кроків із розробки та впровадження правових стандартів такого застосування технологій ШІ, вони лише побіжно вказують на необхідність «виробити єдині принципи і правила їх використання»¹⁹.

Щодо міжнародних стандартів, то Європейська комісія з питань ефективності правосуддя (CEPEJ) на своєму 31-му пленарному засіданні 3–4 грудня 2018 року ухвалила Європейську етичну хартію щодо використання штучного інтелекту у судочинстві. У ній визначено п'ять принципів, якими слід керуватися при впровадженні та застосуванні технологій ШІ у сфері судочинства, які ми розглянемо нижче²⁰.

Цей документ є не єдиною спробою окреслити певні моральні підходи та принципи щодо використання штучного інтелекту у судочинстві. У хартії визначено п'ять етичних принципів, яким має відповідати таке використання, а саме:

1. принцип поваги до основних прав людини: сумісність розробки та впровадження інструментів ШІ з основними правами людини;
2. принцип недискримінації: запобігання будь-якій дискримінації окремих осіб чи груп осіб;

16 American Bar Association House of Delegates Resolution №112 (2019). URL : <https://www.americanbar.org/content/dam/aba/directories/policy/annual-2019/112-annual-2019.pdf> (date requested : 12.01.2022)

17 France Bans Analytics of Judges' Decisions. URL : <https://www.lexology.com/library/detail.aspx?g=ff53dfbe-0fe6-4dee-8a1d-990bf8459020> (date requested : 12.01.2022)

18 Constitutional Council Annual Report 2019. Constitutional Council: Paris 2019. Pp. 42-43.

19 Про схвалення Концепції розвитку штучного інтелекту в Україні: Розпорядження Кабінету Міністрів України від 2 грудня 2020 р. № 1556-р. Урядовий кур'єр. 18.12.2020. №247.

20 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. Council of Europe; European Commission for the Efficiency of Justice (CEPEJ): Strasbourg 2018. Pp. 7-13.

3. принцип якості та безпеки: при обробці судових рішень і даних слід використовувати лише перевірені джерела та дані у спосіб, розроблений згідно з мультидисциплінарним підходом та у захищеному технологічному середовищі;
4. принцип прозорості, неупередженості та справедливості: доступність і зрозумілість методів обробки даних, допустимість їх зовнішніх аудитів;
5. принцип «під контролем користувача»: виключення директивного підходу і гарантування, що користувачі є поінформованими та вільними у своєму виборі використання чи відмови від використання технологій ШІ²¹.

Деякі з цих принципів сформульовані на основі ідей, які вже висловлювалися з приводу практичного застосування штучного інтелекту. Наприклад, принцип прозорості відповідає т.зв. «вимозі чорної скриньки» (black box requirement), за аналогією з пристроєм, який використовують у літаках. У рекомендації Європейського Парламенту від 16 лютого 2017 року її визначено як програмний механізм, «що фіксує дані про кожну транзакцію, вчинену машиною, включно з логікою, яка сприяла прийняттю нею рішень»²². Такий інструмент може допомогти встановити осіб, відповідальних за помилки чи порушення при застосуванні технологій ШІ, і попередити їх у майбутньому.

Це лише перші спроби встановити етичні засади, яким мали би підпорядковуватися розробка та застосування технологій ШІ в правосудді, та сформулювати певне бачення тих основ, на яких повинні ґрунтуватися ці процеси.

21 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. Council of Europe ; European Commission for the Efficiency of Justice (CEPEJ) : Strasbourg, 2018. P. 10.

22 Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) URL: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html (date requested : 12.01.2022)

Світоглядні виклики застосування ШІ в правосудді

Перспективи та наслідки застосування ШІ обговорюються практично у всіх сферах суспільного життя: від маркетингу і сільського господарства до забезпечення громадського порядку і військових озброєнь. Одночасно активно дискутують ризики цього застосування, зокрема економічні (зростання безробіття через зникнення чи трансформацію окремих професій), соціальні (поглиблення нерівності між тими, хто володітиме і користуватиметься благами цих технологій, і тими, хто не матиме такої змоги), ризики пов'язані з правами людини (право на приватність та ін.).

Однак ризики, пов'язані із застосуванням технологій ШІ у публічній сфері є дуже високими. Останніми роками дедалі активнішими стають дискусії про застосування технологій ШІ в процесі прийняття публічних рішень органами влади та правосуддя. Більше того: суспільства з низьким рівнем верховенства права та довіри до судової системи (як-от Україна) розглядають його як простий спосіб вийти з цієї кризи.

Проте бездумне впровадження технологій ШІ на основі суто технічних завдань криє у собі ризики дискримінації (при автоматизованому прийнятті рішень), дегуманізації процесу правосуддя, ризик увійти в суперечність із базовими методами правосуддя (виправлення як мета покарання, врахування обставин справи і поточної суспільної ситуації тощо). Правосуддя може втратити свій сенс і стати інструментом узурпації влади групами, які «програмуватимуть алгоритми».

Аристотель у «Політиці», ведучи мову про правосуддя, називав його виміром справедливості. Прихильники застосування ШІ в правосудді часто не задумуються над тим, що саму ідею справедливості буде дуже складно, якщо взагалі можливо, перевести у машинний код або ж навчити технології (machine learning) розуміти це поняття. А тимчасом саме забезпечення справедливості люди передовсім очікують від правосуддя і саме для цього, зокрема, створена ця гілка влади. І якщо ми звинувачуємо суддів-людей, що вони приймають несправедливі рішення, а отже – якщо відкинути фактор недоброчесності – помилково визначають справедливість, то як ми зможемо навчити цього технології ШІ? З іншого боку – чи будемо ми готові визнавати як справедливі рішення, створені у такий спосіб?

Зрештою, навіть якщо ми відкинемо судові рішення, винесені в минулому, як основу для роботи технологій ШІ в правосудді, а візьмемо для цього закон, то знову опинимося у тій же пастці. Якби мова йшла про бінарну парадигму «добро/зло», то механічне застосування законодавства могло б мати сенс, однак правосуддя значно складніше, і найкращими суддями будуть не ті, які найкраще знають закон, а ті, які його найкраще розуміють.

Тому в дискусіях про перспективи використання технологій ШІ в правозастосуванні, в тому числі у сфері правосуддя, звучить думка, що розробкою правових засад такого використання повинні займатися юристи, які є прихильниками природно-правового підходу. Американський професор права Джошуа П. Дейвіс та філософ Мануель Р. Варґас на прикладі «моральної дилеми автопілота»²³ говорять про те, що юридичні консультації саме від правників із цим підходом будуть найкращими і що у ситуаціях, коли доводиться приймати складне рішення з моральною аргументацією, саме природно-правове бачення «може найкраще пояснити природу права»²⁴.

Можна констатувати, що для належного впровадження ШІ в судочинстві існує потреба у забезпеченні відкритості та прозорості цього процесу, що передбачає, зокрема, залучення широкого кола експертів з різних сфер (мультидисциплінарний підхід), а не лише правників чи ІТ-фахівців.

У Європейській хартії наголошується, що при розробці моделей «машинного навчання» для ШІ у сфері правосуддя необхідно широко використовувати досвід не лише практиків системи правосуддя (суддів, прокурорів, юристів тощо), а й дослідників у галузі права та соціальних наук (економістів, соціологів, філософів)²⁵.

Такий підхід знаходить підтримку і в Європейському Союзі. Так, у рекомендації Комітету з правових питань Європейського Парламенту від 27 січня 2017 року наголошується, що експертні групи повинні бути міждисциплінарними, щоб забезпечувати експертизу з філософських, юридичних та етичних питань, мати у своєму складі членів, які представляють медичну, освітню, соціальну та інші сфери²⁶. Отже, критично важливою для належного впровадження технологій ШІ в правосудді є не лише високопрофесійна технічна та юридична експертиза, а й залучення експертів з інших сфер.

Слід також зазначити, що технології прогнозованого правосуддя ґрунтуються на вже сформованій судовій практиці, в тому числі тих рішеннях, які суспільство критикує і вважає несправедливими. То чи варто покладатися на таку модель при розробці алгоритмів, ризикуючи потрапити у замкнене коло несправедливості, цього разу вже по-справжньому замкнене? Ця проблема підводить нас до ще одного питання: чи не є застосування технологій ШІ в правосудді надто серйозною справою, щоб залишати її лише юристам?

23 Моральна дилема автопілота – проблема вибору поведінки самокерованого автомобіля в ситуації неможливості уникнення дорожньо-транспортної пригоди та неминучого вибору між життям пішохода і водія.

24 Davis J. P. Law Without Mind: AI, Ethics, and Jurisprudence. California Western Law Review. 2018. Vol. 44, No. 1. Pp. 172, 196.

25 European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. Council of Europe ; European Commission for the Efficiency of Justice (CEPEJ) : Strasbourg, 2018. P. 10.

26 Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) URL: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html (date requested : 12.01.2022)

Залучення міждисциплінарної експертизи допоможе забезпечити належні моральні та етичні засади для впровадження змін, що їх принесуть технології ШІ в правосуддя, та баланс інтересів суспільства в цьому плані. Важливий внесок у цей процес можуть зробити представники релігійних конфесій, зокрема Католицької Церкви.

На використання технологій ШІ в правосудді дуже чутливо реагують релігійні спільноти: наприклад, уже лунають застереження, що розвиток технологій ШІ може мати вплив на окремі постулати християнського віровчення. Наприклад, американський християнський автор Джонатан Мерітт (Jonathan Merritt) називає ШІ «найбільшим викликом для християнського богослов'я з часів публікації "Походження видів" Чарльза Дарвіна». З іншого боку, Кевін Келлі (Kevin Kelly), практикуючий католик і редактор-засновник технологічного журналу «Wired», вказує, що «створення речей, які самостійно мислять, можуть викликати серйозні релігійні протиріччя»²⁷.

У соціальному вченні Католицької Церкви сформульовано принципи, які стосуються саме здійснення правосуддя. Наприклад, папа Пій XII у своєму вітальному слові до учасників Першого національного конгресу Союзу італійських католицьких юристів виділив чотири основні стандарти діяльності правників:

1. суддя, виносячи вирок, не може просто перекладати відповідальність на закон та законодавця – він є співвідповідальним разом з правотворцями як за тлумачення змісту закону, так і за наслідки його застосування;
2. суддя своїм рішенням не може змушувати будь-яку особу діяти явно неморально, в тому числі всупереч Закону Божому чи Церкві;
3. суддя не повинен визнавати і схвалювати несправедливий закон (який, по суті, ніколи не може бути по совісті та перед Богом основою чинного вироку), тому не повинен виносити вирок, який був би еквівалентний схваленню такого закону. А якщо таке рішення спричинить публічний скандал, то його відповідальність буде ще більшою;
4. однак не кожне застосування несправедливого закону можна вважати його схваленням, зокрема тоді, коли його застосування буде засобом запобігти більшому злу або отримати добрі наслідки, які матимуть набагато більше значення²⁸.

Аналіз цих стандартів дає можливість зробити певні висновки щодо допустимості застосування ШІ в правосудді. Навіть їх поверховий аналіз дає

27 Is AI a Threat to Christianity? URL : <https://www.theatlantic.com/technology/archive/2017/02/artificial-intelligence-christianity/515463/> (date requested : 12.01.2022)

28 His Holiness Pope Pius XII's Address to the Delegates of the First National Congress of Italian Lawyers (November 6, 1949). URL : <https://aguernz.medium.com/addresses-of-pius-xii-on-purposes-of-punishment-and-the-death-penalty-47ced4c9130e#a4f0> (date requested : 12.01.2022)

підстави стверджувати, що суддя, окрім академічних знань, практичного досвіду та розуміння прецедентного права, має володіти широким розсудом, глибоким розумінням категорій професійної етики і справедливості та, що найважливіше, волею до прийняття складних рішень та усвідомлювати свою відповідальність за них.

Офіційна позиція Католицької Церкви щодо ШІ загалом, так і щодо їх застосування у правосудді зокрема, ще формується. Наприклад, папа Франциск у листопаді 2020 року молився про те, «щоб прогрес робототехніки та штучного інтелекту завжди служив людству»²⁹. Така позиція в цілому відповідає загальноприйнятому у поточному правовому регулюванні підходу – ШІ повинен доповнювати, а не замінювати людину (*AI should complement, but not replace human capabilities*)³⁰.

Папська академія «За життя» організувала у Римі 28 лютого 2020 року конгрес «RenAIssance. Гуманоцентричний штучний інтелект», за підсумками якого було прийнято «Римський заклик до етики штучного інтелекту» (*Rome Call For AI Ethics*). Цей документ отримав підтримку папи Франциска, а також був підписаний такими ІТ-гігантами, як «Майкрософт» та IBM, а також Міжнародною організацією з продовольства та Міністерством інновацій Італії³¹.

Однією з трьох цільових сфер цього документа (англ. impact areas) було визначено сферу прав (англ. rights) та зазначено потребу у «правилах і принципах, які захищають людей, особливо слабких і знедолених, та природне середовище»³². Документ, розроблений з ініціативи Ватикану, пропонує при розробці й застосуванні ШІ дотримуватися шістьох етичних принципів:

- 1) прозорості: має бути зрозуміло, як працюють системи ШІ;
- 2) інклюзивності: мають бути враховані потреби всіх осіб, щоб кожен міг скористатися вигодами, та усім можна було запропонувати найкращі умови для самовираження й розвитку через застосування ШІ;
- 3) відповідальності: ті, хто розробляє та впроваджує використання ШІ, повинні діяти відповідально та відкрито;
- 4) неупередженості: не розробляти і не використовувати ШІ упереджено, захищаючи цим справедливість і гідність людини;
- 5) надійності: системи ШІ повинні бути надійними;

29 Pope's November prayer intention: that progress in robotics and AI "be human". URL : <https://www.vaticannews.va/en/pope/news/2020-11/pope-francis-november-prayer-intention-robotics-ai-human.html> (date requested : 12.01.2022)

30

31 The Pope's plan to fight back against evil AI. URL : <https://www.vox.com/recode/2020/2/28/21157760/pope-vatican-artificial-intelligence> (date requested : 21.01.2022)

32 Rome Call for AI Ethics. URL : https://www.romecall.org/wp-content/uploads/2021/02/AI-Rome-Call-x-firma_DEF_DEF_con-firme_.pdf (21.01.2021)

б) безпеки та конфіденційності: системи ШІ повинні працювати безпечно і з повагою до приватності користувачів³³.

Як бачимо, попри свою ширшу предметну сферу (застосування ШІ в цілому, а не лише в правосудді), багато ідей, що містяться в Заклику, суголосні з тими принципами, які закріплені в Європейській етичній хартії. Зокрема, це принципи прозорості та відповідальності; неупередженості і недискримінації; надійності; безпеки та конфіденційності тощо.

Тому принципи Римського заклику можна розглянути на предмет їх застосування у сфері правосуддя. Принцип неупередженості може означати недопустимість дискримінації у судових рішеннях, прийнятих у такий спосіб, а це передбачає особливу увагу до розробки судових технологій ШІ. Принципи інклюзивності та відповідальності теж стосуються розробки і можуть тлумачитися як потреба в залученні широкого кола акторів до цього процесу та вимога до якості матеріалів, методів і даних, які використовуються у процесі навчання цих систем. Принцип безпеки й конфіденційності передбачає як повагу до права на приватність, так і забезпечення контролю над застосуванням ШІ.

Таким чином, ті підходи до ШІ, які містяться у суспільному вченні Католицької Церкви, можуть мати цілком прикладне значення для застосування інструментів ШІ в правосудді. Це свідчить про потенціал релігійних та інших спільнот у справі розробки візії моральних та етичних засад для розв'язання цього питання.

33 Rome Call for AI Ethics. URL : https://www.romecall.org/wp-content/uploads/2021/02/AI-Rome-Call-x-firma_DEF_DEF_con-firme_.pdf (21.01.2021)

Можливі рішення

Не кожна технологічна новація в правосудді створює ризики. Автоматизація рутинних аспектів процесу судочинства відкриває більше можливостей для всіх його учасників при мінімальних ризиках. Технології ШІ можуть розв'язати проблеми довготривалості судових справ (наприклад, в частині обробки інформації), забезпечення доступу до правосуддя (наприклад, генерування процесуальних документів та їх надсилання), вирішення нескладних категорій справ (наприклад, щодо порушень правил дорожнього руху та ін.).

Однак професія юриста – не тільки судді, а й адвоката, прокурора та ін. – як жодна інша, пов'язана з прийняттям рішень. Найбільші побоювання викликає саме ймовірна передача функцій прийняття судових рішень і, як наслідок, визначення векторів розвитку верховенства права від людей до технологій. Ясна річ, що впровадження технологій неминуче, але необхідно, щоб цей процес опирався на зрозумілі моральні та світоглядні засади.

Чи не найбільшу загрозу застосування технологій ШІ в правосудді несе цілісному людському розвитку та людській гідності. Залишається не розв'язаним складне питання, наскільки така ситуація відповідатиме ідеї справедливості та природі правосуддя, зокрема з морально-етичної та богословської точки зору. Застосування ШІ в правосудді створює цілу низку ризиків, для попередження яких необхідно вживати заходів, тоді як бездумне впровадження цих технологій може спровокувати моральну кризу в сфері правосуддя – мабуть, найбільшу з часу виникнення самого інституту судочинства.

Формування моральних та етичних засад, на яких мало б опиратися використання технологій ШІ у правосудді, вже розпочалося. Можна вести мову про те, що вже окреслені певні рамки для розуміння змісту цих засад як у середовищі правників, так і в інших середовищах. Дуже важливими є ті ідеї та концепції, що їх висувають релігійні середовища, зокрема Католицька Церква. Впровадження технологій ШІ у сфері правосуддя, базоване на цих засадах, матиме тверду моральну основу як для їх застосування, так і для визнання отриманих при цьому результатів.

Такими засадами є:

1. прозорість: процеси розробки та впровадження інструментів ШІ в правосуддя мають бути прозорими та відкритими, що гарантує їх невикористання у протиправних цілях та сформує довіру до них;
2. відкритість: до процесу розробки та впровадження технологій ШІ мають бути залучені експерти з різних сфер, що дозволить забезпечити баланс інтересів суспільства та розуміння поняття «справедливості» із різних точок зору;

3. справедливість: розробка технологій ШІ повинна попереджувати прояви дискримінації, приниження людської гідності, порушення прав і свобод людини та ґрунтуватися на даних, які сприятимуть цій меті;
4. безпека: інструменти ШІ в правосудді мають бути безпечними і має існувати можливість зрозуміти логіку прийняття рішень за їх допомогою.

Ті технології слабого ШІ у формі прогнозованого правосуддя, які існують зараз, не здатні комплексно відповідати цим засадам. Відповідно, є потреба продовжити їх обговорення із подальшим закріпленням як на рівні національного законодавства, так і у формі міжнародних документів. Це дозволить мати об'єктивні стандарти для комплексної оцінки тих чи інших інструментів ШІ для усіх учасників процесу правосуддя.

Дуже важливо продовжувати обговорення моральних та етичних принципів застосування ШІ у різних суспільних сферах, зокрема у релігійній. Приклад Католицької Церкви свідчить, що релігійні спільноти можуть долучатися до обговорення проблеми впливу технологій на суспільство вже на дуже ранніх її стадіях. Своєю участю вони захищають не лише вірних, а й різні маргіналізовані групи, які не мають можливості бути представленими в дискурсі, а тим самим попереджують перетворення інструментів ШІ на джерело проблем замість того, щоб допомагати їх розв'язанню.

Ідею створення ШІ спочатку порівнювали зі створенням людини Богом. Проте Господь, творячи «за своїм образом і подобою», усе ж залишився вищою владою відносно людини. Людство, прагнучи створити розум, подібний своєму, готове ділитися і віддавати цьому «інтелекту» владу над собою і своїми стосунками, зокрема у вимірі справедливості – чи не найбільш визначальному для існування й розвитку людського суспільства. Та якщо самі люди не змогли ефективно забезпечити справедливість між собою, чи варто цього очікувати від машинного коду?

Проте, якщо спроби застосування ШІ уже на порядку денному, то моральні та етичні засади таких спроб як ніколи важливо визначити ДО початку застосування ШІ, а не за його підсумками. І це завдання не для окремих суспільних груп чи акторів, а для суспільства в цілому, де кожен голос з цього приводу повинен бути почутим.

Висновки і рекомендації

Коли ми ведемо мову про розробку і розвиток проєктів щодо технологічних інновацій у правосудді, зокрема із використанням технологій ШІ у сфері правосуддя, то слід передовсім брати до уваги стратегічне значення цієї сфери для суспільства. Технології ШІ, доступні нині чи в найближчому майбутньому, доцільно розглядати лише як допоміжні інструменти.

Ширшому застосуванню технологій ШІ у сфері правосуддя повинен передувати пошук ефективних відповідей на пов'язані з цим ризики. Ключовими серед них є:

- можливе обмеження незалежності правосуддя, зокрема посилення контролю за суддями, впливу на їхні рішення, обмеження свобод та приниження гідності людини та ін. Цей ризик особливо значний в умовах авторитарних суспільств, де політичні режими прагнуть інструменталізувати правосуддя для забезпечення власних інтересів та придушення індивідуальних прав людини;
- недосконалість наявних технологій ШІ може призвести до дискримінації при прийнятті рішень та неврахування обставин справ через дегуманізацію правосуддя;
- складність відображення поняття справедливості за допомогою коду у випадку слабого ШІ; а у випадку сильного ШІ (якщо / коли він буде створений) – зміст цього поняття, за будь-яких обставин, буде інший, ніж для людського суспільства та з позицій людської гідності;
- брак широкої моральної, етичної та світоглядної основи для застосування технологій ШІ у правосудді, а також для їхнього використання з позицій людської гідності, професійної етики і релігійної доктрини. Це може викликати недовіру до них та їхнього використання та кризу довіри до правосуддя, яке має ключову роль при формуванні суспільства на засадах рівності та справедливості.

Щоб попередити ці ризики, варто зробити наступні кроки:

1) прийняти на законодавчому рівні політичні рішення щодо розробки та застосування технологій ШІ в правосудді, які б:

- а) фіксували принципи, згідно з якими має розвиватися цей процес,
- б) встановлювали вимоги до відповідних розробок,
- в) передбачали обмеження та заходи контролю в цій сфері;

2) дотримуватись принципів прозорості та відкритості при здійсненні розробок та впровадженні технологій ШІ в правосудді, що передбачає:

а) залучення до цього процесу широкого кола акторів, а учасниками команд, які братимуть участь у цьому процесі, мають бути не тільки IT-фахівці, а й правники, філософи, соціологи, богослови та ін.,

б) широке суспільне обговорення та перехідний період перед початком застосування технологій ШІ;

3) формально закріпити правові, етичні та інші засади для використання інструментів ШІ в правосудді, що дасть змогу отримати об'єктивний стандарт для їх оцінки;

4) напрацювати не тільки технічні, а й юридичні, моральні та етичні засади впровадження та застосування технологій ШІ в правосудді, з метою формування довіри до них та забезпечення їх внеску у вирішення (а не поглиблення!) проблем у сфері правосуддя.