

Мамаєв Ілля Олександрович

*Молодший науковий співробітник НДІ правового забезпечення
інноваційного розвитку НАПрН України*

ORCID: 0000-0001-7683-0180

ПИТАННЯ ОБРОБКИ ДАНИХ В ЮРИДИЧНИХ ТЕХНОЛОГІЯХ, ЗАСНОВАНИХ НА ШТУЧНОМУ ІНТЕЛЕКТІ

Робота надає загальний огляд того, як юридичні технології (Legal Tech), засновані на штучному інтелекті, можуть використовуватися в юридичній галузі та які питання обробки даних при цьому виникають. Для викладення окресленої теми розкриваються деякі базові вимоги, що висуваються General Data Protection Regulation до обробки даних, а також пояснюється загальне розуміння штучного інтелекту та його роботи.

Ключові слова: захист персональних даних, юридичні технології, штучний інтелект, псевдоанонімізація, GDPR.

Mamaiev Illia

*Research Assistant of Scientific and Research Institute of Providing
Legal Framework for the Innovative Development of NALS of Ukraine*

ORCID: 0000-0001-7683-0180

ISSUES OF DATA PROCESSING IN LEGAL TECH BASED ON ARTIFICIAL INTELLIGENCE

Scientific work provides a general overview of how Legal Tech, based on artificial intelligence, can be used in the legal field and what data processing issues arise. Some basic requirements of General Data Protection Regulation to data processing are revealed to explain the topic. The overall understanding of artificial intelligence and its work is also explained.

Keywords: Personal Data Protection, Legal Tech, Artificial Intelligence, GDPR.

Юридичні технології, також відомі як Legal Tech, передбачають використання цифрових технологій та програмного забезпечення для підвищення ефективності надання юридичних послуг та оптимізації процесів в юридичній галузі. Тлумачення цього терміну є доволі широким та може

включати у себе великий спектр технологій: від додатків, які дозволяють юристам та клієнтам знаходити один одного, до аналітичних систем на основі штучного інтелекту (ШІ), які допомагають здійснювати правові дослідження чи практичну юридичну діяльність.

Разом з тим чимало юридичних технологій передбачають роботу з персональними даними, а тому важливим питанням є забезпечення їх надійного захисту. Особливо актуальним цей аспект стає у зв'язку з необхідністю дотримання General Data Protection Regulation (GDPR), який має виконуватися усіма суб'єктами, що бажають працювати на ринку ЄС чи обробляти персональних дані європейських громадян. Таким чином, захищеність персональних даних є не лише етичним та репутаційним питанням, а й передумовою для розширення компанії у європейському напрямку. Розробникам та користувачам юридичних технологій слід пам'ятати про те, що за порушення вимог GDPR максимальний штраф може складати до 4% річного глобального доходу або 20 мільйонів євро, залежно від того, що виявиться більшим. А порушення окремих норм може призвести й до притягнення до кримінальної відповідальності.

Загалом можна виокремити сім основних вимог до захисту даних, які висуваються GDPR: 1) Псевдоанонімізація – унеможливлення встановлення зв'язку інформації з конкретними особами без додаткових джерел; 2) Збереження у шифрованому вигляді – використання актуальних протоколів шифрування та шляхів передачі даних; 3) Недоступність для осіб, яким не надано відповідних прав – чіткий контроль за доступом до конфіденційної інформації, ведення журналів доступу та блокування несанкціонованих спроб отримати інформацію; 4) Цілісність – інформація має бути точною та зберігатись у повному розмірі, доступ відповідних осіб має бути легким та у повному обсязі. 5) Стабільність засобів обробки інформації – сервіси для обробки інформації мають працювати з мінімальною вірогідністю програмних збоїв або некоректної поведінки. 6) Стійкість до несправностей – значний робочий ресурс обладнання та можливість швидкого відновлення дієздатності

систем (зазвичай рекомендується резервування носіїв інформації та дублювання важливих вузлів). 7) Регулярні ТО обладнання та тестування сервісів – поточні перевірки стану та ефективності систем обробки інформації [1]. Варто розуміти, що це далеко не всі вимоги GDPR, однак їх можна вважати основними умовами, дотримання яких забезпечує захист даних (хоча вони, наприклад, не згадають про обов'язок інформування клієнта про збір даних).

Характерним є те, що GDPR не висуває конкретних вимог до засобів, які мають виконуватися для дотримання окреслених умов (наприклад до того, який тип шифрування має використовуватися), а зосереджується на необхідному результаті, оскільки конкретні технологічні рішення можуть виявитися мінливими та швидко втратити актуальність.

Не є виключенням з цих вимог й технології, засновані на штучному інтелекті. Й, хоча для деяких це поняття досі залишається абстрактним та незрозумілим, такі технології все активніше проникають у повсякденне життя, зокрема й юридичну сферу. Однак для того, щоб прояснити, про що мова піде далі, варто окреслити розуміння штучного інтелекту.

По-перше, під «юридичними технологіями на основі ШІ» не слід розуміти «штучних юристів», що замінюють професіоналів. Більш влучним є порівняння штучного інтелекту з асистентом-дослідником, що може проаналізувати величезний масив даних й виявити певні закономірності. По-друге, за своєю сутністю технології штучного інтелекту передбачають «навчання» комп'ютерів міркувати, спілкуватися, навчатися та приймати рішення, кінцевою метою чого є забезпечення виконання певної інтелектуальної роботи, що в іншому випадку виконувалася б людиною. Важливим у цьому аспекті є те, що замість ручного написання правил того, як комп'ютер має інтерпретувати набір даних, використовуються алгоритми машинного навчання, які дозволяють комп'ютеру самостійно визначати правила та приймати рішення [2].

Фундаментальними для такого навчання є три основні процеси роботи когнітивних обчислень: 1) збір інформації; 2) аналіз і спроба зрозуміти інформацію; 3) прийняття рішень на основі цього розуміння. При цьому, як юрист, так й штучний інтелект, стають краще, коли багаторазово повторюють цю процедуру та вчаться виявляти певні закономірності. Насправді, саме рутинна праця та повтори можуть складати більшу частину праці юриста (так, наприклад, зазвичай не має сенсу писати певний договір з нуля, коли можна використати готові частини зі вже укладених договорів такого роду). При цьому ШІ є найбільш корисним саме тоді, коли може виявити певні паттерни та стандартні часто повторювані запити, на які потрібно відповісти, для чого він має бути забезпечений значними наборами даних для навчання, які, ймовірно, містять ці відповіді.

Сфера застосування ШІ в юридичній практиці може охоплювати юридичні дослідження, побудову стратегії судових спорів, юридичні онлайн-послуги для самопомоги, створення моделей вирішення спорів, огляд і аналіз контрактів та інше. Так, наприклад, штучний інтелект може відповісти на конкретне юридичне запитання («Який термін позовної давності щодо X у державі Y?»), якщо навчався на нормативно-правових матеріалах, що мають цю відповідь. Може допомогти в побудові стратегії судового розгляду та виявити певні закономірності на основі аналізу судових реєстрів («Які шанси на успіх цього клопотання перед цим суддею?»). Може виявити ризики, які криються в договорі, проаналізувавши тисячі відповідних договорів, до яких йому надали доступ [3].

В усіх цих випадках штучний інтелект може бути дійсно ефективним помічником (хоча й з певною ймовірністю неправильної інтерпретації даних), що дозволяє позбавити юристів значної частини рутинних справ та дозволити професіоналам сконцентруватися на тому, що не здатен зробити ШІ. Разом з тим усі описані випадки передбачають навчання штучного інтелекту на основі великого масиву юридичних документів, що може нести ряд ризиків, які потрібно враховувати.

По-перше, юридичні документи можуть містити персональні дані (ідентифікація сторін, фінансові дані, коментарі в контексті переговорів тощо). Індексція юридичних документів і можливість пошуку вмісту неминуче передбачає обробку персональних даних, згаданих у цих юридичних документах, незалежно від мети. Вирішенням цієї проблеми може бути псевдоанонімізація даних, як того вимагає GDPR, однак в ручному режимі вона є тривалим процесом, а автоматична (спеціальним програмним забезпеченням) може пропустити неочевидні місця. У зв'язку з цим раціональною може бути відмова від використання документів, що містять багато персональних даних, але мають низьку корисність для навчання ШІ. По-друге, навіть позбавлені персональних даних документи все ще можуть містити конфіденційні дані, зокрема фінансову, комерційну та технічну інформацію (наприклад, ціни, стратегія продажів, система безпеки). Вирішенням може також слугувати ручне видалення цих даних або відмова від використання таких документів. Однак очевидно, що першочергово слід перевірити, чи поширюються на певні юридичні документи певні зобов'язання щодо конфіденційності. По-третє, при використанні юридичних документів для навчання ШІ слід пам'ятати про контекст цих документів. Так, наприклад, старі документи можуть спиратися на недіючі норми законодавства; певні положення контракту можуть бути результатом складних переговорів та компромісів, які недоцільно використовувати в типовому договорі; а юридичний висновок про правомірність звільнення працівника може відрізнятись при зміні певних фактичних обставин [4].

Висновки. Можна зробити припущення, що юридичні технології, засновані на штучному інтелекті, будуть розвиватися та все більше проникати у юридичну сферу, як й в інші сфери життя. У зв'язку з цим раннє освоєння та розвиток технологій ШІ може бути передумовою майбутньої конкурентоспроможності, оскільки потенційно дозволяє звільнити юристів від рутинної повторюваної роботи та оптимізувати робочі процеси. Втім при використанні та розборці юридичних технологій, заснованих на штучному

інтелекті, варто пам'ятати про вимоги щодо захисту даних, зокрема положення GDPR. До таких базових вимог, що забезпечують захист даних, можна віднести: 1) псевдоанонімізацію; 2) збереження у шифрованому вигляді; 3) недоступність для осіб, яким не надано відповідних прав; 4) цілісність; 5) стабільність засобів обробки інформації; 6) стійкість до несправностей; 7) регулярні ТО обладнання та тестування сервісів. В контексті штучного інтелекту особливо актуальною є псевдоанонімізація даних, оскільки процес навчання ШІ передбачає обробку великих масивів юридичних документів, що можуть мати персональні чи конфіденційні дані. Окрім того, розробникам юридичних технологій на основі штучного інтелекту має сенс звертати увагу й на контекст документів (передумови їх прийняття, дійсність тощо), щоб уникнути проблем з навчанням ШІ.

ЛІТЕРАТУРА

1. Тамашев В. Технічні вимоги GDPR. Реалізація на практиці. Legal IT Group website. URL: <https://legalitgroup.com/gdpr-tech-vimogi/>.
2. Miller S. Artificial intelligence and its impact on legal technology: to boldly go where no legal department has gone before. Thomson Reuters website. URL: <https://legal.thomsonreuters.com/en/insights/articles/ai-and-its-impact-on-legal-technology>.
3. Demystifying Artificial Intelligence (AI). Thomson Reuters website. URL: <https://legal.thomsonreuters.com/en/insights/white-papers/demystifying-ai>.
4. Fiten B., Somers G. AI-based legal tech solutions: discover the legal pitfalls. Timelex website, 20.02.2023. URL: <https://www.timelex.eu/en/blog/ai-based-legal-tech-solutions-discover-legal-pitfalls>.

REFERENCES

1. Tamashev, V. GDPR technical requirements. Implementation in practice. Legal IT Group website. Retrieved from: <https://legalitgroup.com/gdpr-tech-vimogi/> [in Ukrainian].

2. Miller, S. Artificial intelligence and its impact on legal technology: to boldly go where no legal department has gone before. Thomson Reuters website. Retrieved from: <https://legal.thomsonreuters.com/en/insights/articles/ai-and-its-impact-on-legal-technology>.

3. Demystifying Artificial Intelligence (AI). Thomson Reuters website. Retrieved from: <https://legal.thomsonreuters.com/en/insights/white-papers/demystifying-ai>.

4. Fiten, B. & Somers, G. (2023). AI-based legal tech solutions: discover the legal pitfalls. Timelex website. Retrieved from: <https://www.timelex.eu/en/blog/ai-based-legal-tech-solutions-discover-legal-pitfalls>.