Ірха Ю.Б.,
кандидат юридичних наук, доцент, заслужений юрист України, завідувач наукової лабораторії права національної та міжнародної безпеки Державної наукової установи «Інститут інформації, безпеки і права Національної академії правових наук України»
(м. Київ, Україна),
доцент кафедри права Дніпровського гуманітарного університету
(м. Дніпро, Україна)

ШТУЧНИЙ ІНТЕЛЕКТ У СФЕРІ НАЦІОНАЛЬНОЇ БЕЗПЕКИ ТА ОБОРОНИ: ПРАВА ЛЮДИНИ, ЛЮДСЬКИЙ КОНТРОЛЬ ТА КУЛЬТУРА ВІДПОВІДАЛЬНОСТІ

Штучний інтелект (далі – ШІ) вже давно перестав бути плодом уявлення письменників-фантастів та продуктом закритих наукових лабораторій. Його відкрите використання та вільний доступ кардинально змінили життя людини, яка користується продуктами сучасного науково-технічного прогресу: Інтернетом, смартфонами, соціальними мережами, навігаційними системами, електромобілями, онлайн-банкінгом, системами розумного дому тощо.

Здатність цифрового розуму швидко обробляти, систематизувати та аналізувати великі обсяги даних; виявляти приховані закономірності; робити прогнози; створювати нову інформацію; автономно діяти та взаємодіяти з реальним та цифровим середовищем для досягнення поставленої мети; самостійно вчитися та удосконалюватися – відкрила величезні можливості у багатьох сферах суспільного життя та істотно прискорила розвиток науки і техніки. Ми отримали надпотужний інструмент для прогресу, однак його потенціал та спроможності дедалі більше викликають побоювань про те, що він зможе не лише слідувати людським інструкціям, але й визначати власні пріоритети, цілі та цінності, які можуть завдати непоправної шкоди людству.

За твердженнями українських експертів, на відміну від звичайних обчислювальних систем, у випадку ШІ спостерігається ефект непередбачуваності (частково – не тривіальності) результатів його «роздумів», що в загальному випадку є однією з ознак творчості та інноваційної діяльності, яка притаманна людині. З іншого боку, відсутність прозорих методів перевірки запропонованих

ШІ висновків та рекомендацій утворює джерело невизначеності щодо їх вірності і практичної цінності [1, с. 7].

У 2023 році засновник SpaceX Ілон Маск, співзасновник Apple Стів Возняк та інші відомі фахівці та вчені у сфері новітніх технологій підписали відкритий лист про призупинення розвитку ШІ. На їхнє переконання, лабораторії ШІ вступили в неконтрольовану гонку з розробки та впровадження все більш потужних цифрових розумів, які ніхто – навіть їхні творці – не може зрозуміти, передбачити або надійно контролювати. Вони застерегли світ від небезпечної гонитви за все більшими непередбачуваними моделями «чорного ящика» з новітніми можливостями.

Автори листа вважають, що дослідження та розробки в галузі ШІ повинні бути переорієнтовані на те, щоб зробити сучасні високотехнологічні системи більш точними, безпечними, зрозумілими, прозорими, стійкими, збалансованими, надійними та лояльними. Підписанти пропонували встановити щонайменше піврічний мораторій на розробку програм ШІ, які є потужнішими за ChatGPT-4, задля формування універсальних протоколів безпеки для прогресивного проектування та розробки ШІ. Ці протоколи мали б гарантувати, що системи, які їх дотримуються, є безпечними поза розумними сумнівами [2].

Цей відкритий лист хоч і не зупинив розвиток ШІ, але в черговий раз привернув увагу до масштабу проблеми. Він показав, що її потрібно вирішувати комплексно: не лише на рівні розробників і виробників, а й на правовому, етичному та безпековому рівнях.

В Асіломарських принципах роботи зі штучним інтелектом наголошено, що:

- передовий ШІ може спричинити глибокі зміни в історії життя на Землі, тому його впровадження та управління ним повинні плануватися з відповідною ретельністю та ресурсами;
- системи ШІ повинні бути безпечними та надійними протягом усього терміну їх експлуатації, а також піддаватися перевірці, де це доцільно та можливо;
- високоавтономні системи ШІ повинні бути розроблені таким чином, щоб їх цілі та поведінка були узгоджені з людськими цінностями протягом усього терміну їх експлуатації;
- системи ШІ повинні бути спроектовані та експлуатуватися таким чином, щоб бути сумісними з ідеалами людської гідності, прав, свобод та культурного різноманіття;
- люди повинні вибирати, яким чином і чи варто делегувати прийняття рішень системам ШІ для досягнення обраних людьми цілей;

- системи ШІ, розроблені для рекурсивного самовдосконалення або самовідтворення, що може призвести до швидкого підвищення їх якості або кількості, повинні підлягати суворим заходам безпеки та контролю [3].

Спроможності ШІ вже давно вийшли за межі людських можливостей і глибоко інтегрувалися у повсякденне життя. Суб'єкти, які явно чи приховано використовують прогресивні моделі ШІ, отримують значні переваги у бізнесі, управлінні, творчості, науці. Наведене неминуче призводить до формування нового «цифрового розриву» між тими, хто має доступ до цих технологій, і тими, хто залишається позаду. Це також створює загрозу неконтрольованої технологічної гонки, учасники якої нехтують фундаментальними правами і свободами людини, етичними нормами та глобальною безпекою.

У новій технологічній реальності виклики та загрози пов'язані із використанням ШІ найбільш гостро постають у сфері національної безпеки та оборони. Держави з метою захисту свого суверенітету, територіальної цілісності, конституційного ладу щоденно намагаються впровадити системи ШІ не тільки в озброєння, бойову техніку, боєприпаси, але й у військове та цивільне управління, фінансово-економічні процеси, механізми збору та обробки інформації в інтересах розвідки, контррозвідки, кримінального судочинства, контролю за інформаційним полем тощо. При цьому такі системи можуть виконувати як допоміжні функції, так і повністю самостійні завдання.

Впровадження різноманітних моделей ШІ у діяльність військових та правоохоронців істотно змінило механізми забезпечення обороноздатності, протидії злочинності, підтримання публічного порядку та безпеки, а також тактику та стратегії ведення бойових дій, проведення антитерористичних та правоохоронних операцій. Використовуючи ШІ, вони прагнуть отримати здатність діяти на випередження в масштабах, які перевищують людські когнітивні можливості. Крім того, передові технології сприяють підвищенню швидкості прийняття рішень, покращенню точності ураження цілей, зменшенню ризиків для особового складу, оптимізації використання ресурсів, а також забезпечують цілодобове функціонування оборонних, розвідувальних та інформаційно-аналітичних комплексів.

Разом з тим, у гонитві за технологічною перевагою все частіше на другий план почали відходити декілька фундаментальних питань. По-перше, як забезпечити контроль за законністю та дотриманням прав і свобод людини під час його використання та/або застосування. По-друге, як визначити юридичну відповідальність розробників, виробників та користувачів за шкоду, завдану інтелектуальною програмою умисно, з необережності чи випадково. По-третє, наскільки можна довіряти інформації та рішенням, які генерує ШІ.

З огляду на те, що параметри передових систем ШІ у сфері національної безпеки і оборони не є публічними, а про їх реальні можливості можуть не знати і самі розробники, то існують серйозні загрози для створення так званої «цифрової диктатури», коли за благими намірами підвищення ефективності управління державою та захисту демократії створюється система, в якій рішення щодо долі людей приймаються не представниками органів публічної влади, а непідзвітною нікому машиною, цінності, логіку та рішення якої неможливо ані зрозуміти, ані оскаржити.

Для України, яка в умовах повномасштабної російської збройної агресії активно впроваджує новітні технології для перемоги, ці загрози є надзвичайно актуальними.

З одного боку, ШІ є одним із інструментів для отримання технологічної переваги над ворогом та забезпечення ефективного безпекового середовища на територіях, де не ведуться бойові дії. З іншого – висока швидкість впровадження ШІ у сферу національної безпеки та оборони створює передумови для прийняття на озброєння недостатньо перевірених технологій ШІ та допуску до їх експлуатації непідготовлених фахівців, що лише підвищує ризики помилок, зловживань, свавілля та непередбачуваних дій. У поєднанні із відсутністю правового регулювання ШІ в Україні це породжує загрози, за яких інструменти, створені для захисту демократії, можуть бути використані для її демонтажу в майбутньому.

Зазначені загрози вимагають формування якісної системи правового регулювання ШІ, яка б інтегрувала найкращі світові практики та відповідала б безпековим викликам та загрозам, що стоять перед Україною. Вітчизняне правове поле у сфері ШІ має одночасно сформувати сприятливі умови для технологічних інновацій та забезпечити захист таких фундаментальних цінностей як верховенство права, демократія, гідність людини, її права і свободи. Для цього на законодавчому та підзаконному рівнях необхідно якнайшвидше визначити:

– засади функціонування ШІ;
– допустимі сфери та межі застосування ШІ;
– механізми державного та недержавного нагляду, сертифікації та аудиту у сфері ШІ;
– вимоги до кібербезпеки, захисту інформації та даних, що використовуються для навчання ШІ;
– правила маркування контенту, створеного за допомогою ШІ;
– права, обов’язки та відповідальність розробників, виробників, користувачів ШІ;

– особливості використання та застосування ШІ в органах сектору безпеки і оборони, у тому числі в летальних автономних системах озброєння та комплексах збору та обробки Big data;

– визначення рівнів та вимог до взаємодії людини і ШІ залежно від ступеня ризику відповідної системи тощо.

Поряд із законодавчими новаціями в українському суспільстві необхідно впроваджувати та розвивати культуру відповідального використання та застосування ШІ, що має ґрунтуватися на двох ключових елементах: освіті та суспільному діалозі. Така культура не може бути нав'язана згори, вона має стати результатом спільних зусиль держави, бізнесу, освітніх та наукових установ, медіа, громадянського суспільства. Кожен громадянин має розуміти не лише переваги, а й загрози сучасних технологій, адже в умовах гібридної війни це є питанням національної стійкості. Лише шляхом підвищення ШІ-грамотності населення та відкритих дискусій можна виробити спільні етичні орієнтири, які стануть основою для безпечної інтеграції ШІ в життя країни.

З огляду на вищезазначене, розробка збалансованого підходу до ШІ є для України не просто технологічним, а й світоглядним завданням, від вирішення якого залежить майбутнє її демократичного розвитку.

Список використаних джерел:

1. Скіцько О., Складанний П., Ширшов Р., Гуменюк М., Ворохоб М. Загрози та ризики використання штучного інтелекту. *Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка»*, 2023. № 2(22), С. 6–18. DOI: https://doi.org/10.28925/2663-4023.2023.22.618.

2. Pause Giant AI Experiments: An Open Letter. URL: https://futureoflife.org/wp-content/uploads/2023/05/FLI_Pause-Giant-AI-Experiments_An-Open-Letter.pdf

3. Asilomar AI Principles. URL: https://futureoflife.org/open-letter/ai-principles