

Без волі немає контролю

Зброєю на основі штучного інтелекту можна керувати. Чому тоді найбільші перешкоди стосуються не технічного, а політичного аспекту?

Штучний інтелект (ШІ) кардинально змінює не лише сферу праці, а й принципи ведення війни. Якщо зброя на базі ШІ – майбутнє війни, то контроль над озброєнням у цій галузі стає нагальною потребою. Багато хто скаже, що традиційні методи контролю над озброєнням не можна застосувати до зброї на базі ШІ. Дослідження, зокрема, вказують на принципові відмінності між контролем над ядерною зброєю та контролем над зброєю на базі ШІ, підкреслюючи суттєві труднощі у встановленні контролю над останніми.

Попри те, що скептицизм небезпідставний, справжньою проблемою у сфері контролю над озброєннями на базі ШІ є брак політичної волі. У статті я пропоную розглянути такі питання, як можливість перевірки, правові та етичні «червоні лінії», використання технологій подвійного призначення та можливі майбутні інституційні механізми.

По-перше, хтось стверджує, що алгоритми ШІ та набори не піддаються жодній перевірці, бо не мають фізичної форми. Проте навіть у випадку таких договорів, як Новий договір про скорочення стратегічних наступальних озброєнь (New START) між США та Росією, який до свого завершення у лютому 2026 року передбачав проведення інспекцій на місцях і обмін даними про ядерну зброю, багато хто сумнівався, що можна досягти надійної перевірки й запобігти обману з боку іншої сторони. Вони стали підсумком багаторічних переговорів, комунікації та зміцнення довіри.

Крім того, провести повну та прозору перевірку всього арсеналу ядерної зброї просто неможливо. Дві сторони відібрали частину найважливіших стратегічних ядерних озброєнь і запровадили так званий механізм інспекцій з контролю над озброєннями, що ґрунтувався на конкретному обміні даними та частково відкритих інспекціях, які підтримувалися технологіями, що постійно вдосконалювалися. Контроль над ядерною зброєю на основі

верифікації залишався б надзвичайно складним без готовності США та Росії частково розкривати свої ядерні арсенали. Інші держави, що володіють ядерною зброєю, також не допускають таких інспекцій, і це призводить до суперечок.

Упродовж останніх років на впровадження механізмів верифікації контролю над озброєннями між США та Росією негативно вплинули геополітичні конфлікти, такі як війна в Україні. Обидві держави вважають, що інша сторона активно просуває модернізацію свого ядерного арсеналу. А це, в свою чергу, свідчить про те, що верифікація ядерної зброї залежить від певних передумов, зокрема від політичного посередництва. Через те, що об'єкт контролю має фізичну форму і його можливо порохувати, не означає, що ядерну зброю перевіряти легше.

Перевірка алгоритмів та наборів даних ШІ стикається з низкою технічних перешкод, але це не означає, що їх неможливо перевірити. Насправді немає потреби перевіряти всі моделі та набори даних, адже для тестування й оцінки країни повинні надати лише ті моделі й набори даних, які становлять найбільшу загрозу безпеці. Тестування безпеки ШІ стикається з тими ж викликами, що й *управління* кібербезпекою. Теоретично міжнародна спільнота могла б створити незалежні лабораторії для оцінки прогалин у кібербезпеці й сформувати бази даних щодо вразливостей і шкідливого програмного забезпечення. Незалежні лабораторії можна було б запросити для проведення оцінок кібербезпеки критичної цифрової інфраструктури та важливих ланок постачання. Той факт, що цю ідеальну модель досі не втілили на практиці, пояснюється значним рівнем взаємної недовіри великих держав одна до одної та розбіжностями інтересів у сфері кібербезпеки.

Загалом брак політичної волі до співпраці та довіри є значно більшою перешкодою, ніж технічні проблеми

Загалом брак політичної волі до співпраці та довіри є значно більшою перешкодою, ніж технічні проблеми. Показовим прикладом є ситуація, коли у 2018 році китайська компанія Huawei, прагнучи отримати доступ до британського ринку 5G, погодилася пройти технічну перевірку британськими органами безпеки і створила в Брюсселі центр прозорості кібербезпеки, де надавався б доступ до їхнього вихідного коду й тестових середовищ. Багато європейських держав відмовилися від впровадження технології 5G від Huawei з політичних міркувань.

По-друге, деякі стверджують, що правові та етичні «червоні лінії» щодо ядерної зброї є дуже чіткими, що полегшує досягнення міжнародного консенсусу стосовно контролю над ядерним озброєнням порівняно з регулюванням ШІ. Навіть «червоні лінії» і «табу», пов'язані з ядерною зброєю, залишаються невизначеними, через що виникає багато суперечок. Наприклад, згідно з рішенням Міжнародного суду 1994 року, досі неможливо однозначно визначити, чи є законним застосування ядерної зброї для захисту своєї держави. Договір про нерозповсюдження ядерної зброї (Non-Proliferation Treaty, NPT) не є достатньо точним, коли йдеться про застосування ядерної зброї в інших країнах.

По-третє, хтось стверджує, що ядерна зброя є специфічною технологією, для якої характерне чітке розмежування між військовим і цивільним застосуванням, а тому навколо військових застосувань легко створити окрему систему контролю над озброєннями. Натомість ШІ настільки поширений і застосування його настільки універсальне, що складно провести чітке розмежування між військовими та цивільними сферами використання. Така точка зору не враховує основного: революція в ядерній технології відіграла вирішальну роль у світовому економічному відновленні після Другої світової війни, як це ілюструє програма «Мирний атом». Крім того, системи контролю над озброєннями в ядерній сфері ніколи не обмежувалися лише військовою сферою. Переконавання, що військове та цивільне застосування ядерних технологій легко відрізнити, є результатом півстолітніх спільних зусиль держав, спрямованих на встановлення всеосяжного режиму нерозповсюдження ядерної зброї, що охоплює як військові, так і цивільні сфери використання, а також систем експортного контролю для запобігання поширенню ядерної зброї, таких як Комітет Зангера (Zangger Committee).

Уявлення про те, що в ядерних технологіях легко відрізнити військове використання від цивільного, сформувалося під впливом багаторічної дії Договору про нерозповсюдження ядерної зброї (NPT). Три його засади – нерозповсюдження, ядерне роззброєння та мирне використання ядерної енергії – надійно закріпилися у свідомості людей. Для запобігання розповсюдження ядерної зброї США використовували різні засоби, такі як дипломатичний тиск, економічні санкції, навіть військові заходи. Правила стосовно застосування ядерної зброї дають більшості держав чіткі вказівки щодо того, що слід робити, а від чого утриматися.

Водночас багато країн намагалися отримати ядерну зброю шляхом розвитку цивільної ядерної енергетики. Іранське ядерне питання досі залишається невирішеним, оскільки існують розбіжності щодо

прийняттого рівня збагачення урану. Деякі країни стали так званими ядерними державами-початківцями, розвиваючи потужний цивільний ядерний потенціал та накопичуючи запаси чутливих ядерних матеріалів, тобто таких, що вимагають найвищого рівня захисту через ризик радіоактивних викидів.

Найбільша проблема в тому, що ми поки не дійшли згоди стосовно того, які технології та сфери застосування ІІІ є неприйнятними

Подвійне призначення технологій ІІІ не є головною перешкодою для механізмів контролю над озброєннями на базі ІІІ. Об'єктами таких механізмів майже завжди є технології подвійного призначення. Найбільша проблема в тому, що ми поки не дійшли згоди стосовно того, які технології та сфери застосування ІІІ є неприйнятними. У цьому сенсі ми наче живемо в часи до 1945 року, коли потужні держави вели інтенсивну боротьбу за революційну технологію, проте її остаточний вигляд залишався невідомим.

Перша держава, що отримає Загальний штучний інтелект ЗІІІ (Artificial General Intelligence – AGI) або Штучний надінтелект ШНІ (Superintelligent AI, ASI), чия приголомшливу силу ми ще до кінця не розуміємо, ймовірно, спробує встановити монополію. Коли інші країни поступово отримають AGI або ASI, вони, мабуть, випробують нові механізми нерозповсюдження ядерної зброї та міжнародної безпеки. Можливо, ці інші держави зможуть застосовувати технології ІІІ, однак їм, буде заборонено користуватися AGI або ASI повною мірою, аналогічно до обмежень, встановлених Договором про нерозповсюдження ядерної зброї. Відповідні механізми нерозповсюдження можуть бути створені щодо обчислювальної потужності, наборів даних та ресурсів електроенергії. Великі держави могли б досягти нового рівня взаємної вразливості та стратегічної стабільності шляхом придбання AGI або ASI.

Країни, що розвиваються, та країни Глобального Півдня не схвалять створення ще одного дискримінаційного міжнародного механізму, який збільшить розрив у сфері ІІІ й обмежить їхнє право на розвиток. Альтернативним сценарієм майбутнього інституційного регулювання могло б стати створення регіональних обчислювальних та енергетичних центрів у Східній Азії, Південній Азії, на Близькому Сході, в Європі, Африці та Америці, а також формування кооперативних організацій на базі вже наявних механізмів регіональної інтеграції, які відповідали б за дослідження у сфері ASI та виявлення ризиків для безпеки.



Тяньцзяо Цзян

Тяньцзяо Цзян (Tianjiao Jiang) – доцент Інституту розвитку при Університеті Фудань. З 2026 року – запрошений науковець у Китайському центрі ім. Пола Цая при Юридичній школі Єльського університету. Нові технології та міжнародна безпека – основні напрями його наукових досліджень.

