

Без воли нет контроля

Оружием на основе ИИ можно управлять. Самые большие препятствия связаны не с техническим, а с политическим аспектом.

Искусственный интеллект (ИИ) кардинально меняет не только сферу труда, но и принципы ведения войны. Если оружие на базе ИИ – будущее войны, то контроль над вооружением в этой области становится насущной необходимостью. Многие скажут, что традиционные методы контроля над вооружениями нельзя применить к оружию на базе ИИ. Исследования, в частности, указывают на принципиальные различия между контролем над ядерным оружием и контролем над оружием на базе ИИ, подчеркивая существенные трудности в установлении контроля над последним.

Несмотря на то, что скептицизм небезоснователен, настоящей проблемой в сфере контроля над вооружениями на базе ИИ является отсутствие политической воли. В статье я предлагаю рассмотреть такие вопросы, как возможность проверки, правовые и этические «красные линии», использование технологий двойного назначения и возможные будущие институциональные механизмы.

Во-первых, некоторые утверждают, что алгоритмы ИИ и наборы данных не поддаются никакой проверке, поскольку не имеют физической формы. Однако даже в случае таких договоров, как Новый договор о сокращении стратегических наступательных вооружений (New START) между США и Россией, который до своего завершения в феврале 2026 года предусматривал проведение инспекций на местах и обмен данными о ядерном оружии, многие сомневались, что можно добиться надежной проверки и предотвратить обман от другой стороны. Они стали итогом многолетних переговоров, коммуникации и укрепления доверия.

Кроме того, провести полную и прозрачную проверку всего арсенала ядерного оружия просто невозможно. Обе стороны отобрали часть важнейших стратегических ядерных вооружений и ввели так называемый механизм инспекций по контролю над вооружениями,

основанный на конкретном обмене данными и частично открытых инспекциях, которые поддерживались постоянно совершенствующимися технологиями. Контроль над ядерным оружием на основе верификации оставался бы чрезвычайно сложным без готовности США и России частично раскрывать свои ядерные арсеналы. Другие государства, обладающие ядерным оружием, также не допускают таких инспекций, и это приводит к спорам.

В последние годы на внедрение механизмов верификации контроля над вооружениями между США и Россией негативно повлияли геополитические конфликты, такие как война в Украине. Обе страны считают, что другая сторона активно продвигает модернизацию своего ядерного арсенала. А это, в свою очередь, свидетельствует о том, что верификация ядерного оружия зависит от определенных предпосылок, в частности от политического посредничества. Тот факт, что объект контроля имеет физическую форму и его можно сосчитать, не означает, что ядерное оружие проверять легче.

Проверка алгоритмов и наборов данных ИИ сталкивается с рядом технических препятствий, но это не означает, что их невозможно проверить. На самом деле нет необходимости проверять все модели и наборы данных, ведь для тестирования и оценки страны должны предоставить только те модели и наборы данных, которые представляют наибольшую угрозу безопасности. Тестирование безопасности ИИ сталкивается с теми же вызовами, что и *управление* кибербезопасностью. Теоретически международное сообщество могло бы создать независимые лаборатории для оценки пробелов в кибербезопасности и сформировать базы данных об уязвимостях и вредоносном программном обеспечении. Независимые лаборатории можно было бы пригласить для проведения оценок кибербезопасности критической цифровой инфраструктуры и важных звеньев цепочки поставок. Тот факт, что эта идеальная модель до сих пор не воплощена на практике, объясняется значительным уровнем взаимного недоверия крупных государств друг к другу и расхождениями интересов в сфере кибербезопасности.

В целом отсутствие политической воли к сотрудничеству и доверия является гораздо большим препятствием, чем технические

В целом отсутствие политической воли к сотрудничеству и доверия является гораздо большим препятствием, чем технические проблемы. Показательным примером является ситуация, когда в 2018 году китайская компания Huawei, стремясь получить доступ к британскому рынку 5G, согласилась пройти техническую проверку

британскими органами безопасности и создала в Брюсселе центр прозрачности кибербезопасности, где предоставлялся бы доступ к их исходному коду и тестовым средам. Многие европейские государства отказались от внедрения технологии 5G от Huawei по политическим соображениям.

Во-вторых, некоторые утверждают, что правовые и этические «красные линии» в отношении ядерного оружия очень четкие, что облегчает достижение международного консенсуса по контролю над ядерным вооружением по сравнению с регулированием ИИ. Даже «красные линии» и «табу», связанные с ядерным оружием, остаются неопределенными, из-за чего возникает много споров. Например, согласно решению Международного суда 1994 года, до сих пор невозможно однозначно определить, является ли законным применение ядерного оружия для защиты своего государства. Договор о нераспространении ядерного оружия (Non-Proliferation Treaty, NPT) не является достаточно точным, когда речь идет о применении ядерного оружия в других странах.

В-третьих, некоторые утверждают, что ядерное оружие представляет собой специфическую технологию, для которой характерно четкое разграничение между военным и гражданским применением, и поэтому вокруг военных применений легко создать отдельную систему контроля над вооружениями. В то же время ИИ настолько распространен, а его применение настолько универсально, что сложно провести четкое разграничение между военной и гражданской сферами использования. Такая точка зрения не учитывает главного: революция в ядерной технологии сыграла решающую роль в мировом экономическом восстановлении после Второй мировой войны, как это иллюстрирует программа «Мирный атом». Кроме того, системы контроля над вооружениями в ядерной сфере никогда не ограничивались только военной сферой. Убеждение, что военное и гражданское применение ядерных технологий легко отличить, является результатом полувековых совместных усилий государств, направленных на установление всеобъемлющего режима нераспространения ядерного оружия, охватывающего как военные, так и гражданские сферы использования, а также систем экспортного контроля для предотвращения распространения ядерного оружия, таких как Комитет Зангера (Zangger Committee).

Представление о том, что в ядерных технологиях легко отличить военное применение от гражданского, сформировалось под

влиянием многолетнего действия Договора о нераспространении ядерного оружия (НРТ). Три его принципа – нераспространение, ядерное разоружение и мирное использование ядерной энергии – прочно укоренились в сознании людей. Для предотвращения распространения ядерного оружия США использовали различные средства, такие как дипломатическое давление, экономические санкции и даже военные меры. Правила применения ядерного оружия дают большинству государств четкие указания относительно того, что следует делать, а от чего воздерживаться.

В то же время многие страны пытались получить ядерное оружие путем развития гражданской ядерной энергетики. Иранский ядерный вопрос до сих пор остается нерешенным, поскольку существуют разногласия относительно приемлемого уровня обогащения урана. Некоторые страны стали так называемыми начинающими ядерными державами, развивая мощный гражданский ядерный потенциал и накапливая запасы чувствительных ядерных материалов, то есть таких, которые требуют высочайшего уровня защиты из-за риска радиоактивных выбросов.

Самая большая проблема в том, что мы пока не пришли к согласию относительно того, какие технологии и сферы применения ИИ являются неприемлемыми

Двойное назначение технологий ИИ не является главным препятствием для механизмов контроля над вооружениями на базе ИИ. Объектами таких механизмов почти всегда являются технологии двойного назначения. Самая большая проблема в том, что мы пока не пришли к согласию относительно того, какие технологии и сферы применения ИИ являются неприемлемыми. В этом смысле мы как будто живем в эпоху до 1945 года, когда могущественные государства вели интенсивную борьбу за революционную технологию, но ее окончательный облик оставался неизвестным.

Первое государство, которое получит Общий искусственный интеллект (AGI) или Искусственный сверхинтеллект (ASI), чью ошеломляющую силу мы еще до конца не понимаем, вероятно, попытается установить монополию. Когда другие страны постепенно получают AGI или ASI, они, вероятно, испытают новые механизмы нераспространения ядерного оружия и международной безопасности. Возможно, эти другие государства смогут применять технологии ИИ, но им будет запрещено пользоваться AGI или ASI в полной мере, аналогично ограничениям, установленным Договором

о нераспространении ядерного оружия. Соответствующие механизмы нераспространения могут быть созданы в отношении вычислительной мощности, наборов данных и ресурсов электроэнергии. Крупные государства могли бы достичь нового уровня взаимной уязвимости и стратегической стабильности путем приобретения AGI или ASI.

Развивающиеся страны и страны Глобального Юга не одобряют создание еще одного дискриминационного международного механизма, который усугубит разрыв в сфере ИИ и ограничит их право на развитие. Альтернативным сценарием будущего институционального регулирования могло бы стать создание региональных вычислительных и энергетических центров в Восточной Азии, Южной Азии, на Ближнем Востоке, в Европе, Африке и Америке, а также формирование кооперативных организаций на базе уже существующих механизмов региональной интеграции, которые отвечали бы за исследования в сфере ASI и выявление рисков для безопасности.

По этой ссылке статья откроется без VPN



Тяньцзяо Цзян

Тяньцзяо Цзян (Tianjiao Jiang) – доцент Института развития при Университете Фудань. С 2026 года – приглашенный ученый в Китайском центре им. Пола Цая при Юридической школе Йельского университета. Новые технологии и международная безопасность – основные направления его научных исследований.



