

Документ Российской Федерации во исполнение резолюции ГА ООН 79/239 от 24 декабря 2024 г. «Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности»

Российская Федерация отмечает принятие резолюции ГА ООН 79/239 от 24 декабря 2024 г. и в соответствии с её пунктом 7 имеет честь представить свой национальный вклад к докладу Генерального секретаря ООН 80-й сессии ГА ООН для дальнейшего обсуждения государствами-членами.

Введение

Российская Федерация придает важное значение вопросам применения технологии искусственного интеллекта (ИИ) в военной области. Заинтересованы в дальнейшем предметном обсуждении данной проблематики на профильных международных площадках.

Рассматриваем Группу правительственных экспертов государств-участников Конвенции о «негуманном» оружии (КНО) по смертоносным автономным системам вооружений (ГПЭ по САС) в качестве оптимальной площадки для такой дискуссии. Именно профильная ГПЭ призвана выдерживать разумный баланс между гуманитарными озабоченностями и законными оборонными интересами государств в отношении таких средств, а также принимать решения на основе консенсуса. Рассмотрение вопросов военного применения ИИ в рамках Группы имеет широкий охват, не сводится сугубо к проблематике САС и затрагивает ряд важных аспектов (в т.ч. правового, технического и военного плана), связанных с задействованием данной технологии в военных целях.

Отмечаем обсуждение указанной тематики в рамках действующих режимов в области контроля над вооружениями, разоружения и нераспространения (КВРН). Данная работа сфокусирована на анализе рисков и возможностей, которые предоставляет ИИ в плане выполнения государствами-

участниками обязательств, предусмотренных соответствующими международно-правовыми инструментами.

Приветствуем готовность государств-членов приступить к обсуждению тематики военного применения ИИ в Комиссии ООН по разоружению (КОР) в рамках дискуссии о новых технологиях в контексте международной безопасности. Данный обмен мнениями направлен на согласование рекомендаций по таким аспектам «военного» ИИ, которые не рассматриваются на других площадках.

В ходе работы на упомянутых международных форумах требуют особого внимания вопросы разработки единой терминологической базы, применения действующего международного права, осуществления человеческого контроля, обеспечения ответственности, а также создаваемых данной технологией рисков и возможностей.

Определение

В действующем международном праве отсутствует консенсусное определение систем вооружений и военной техники с использованием ИИ, что осложняет рассмотрение данной проблематики. Выработка единого рабочего понимания подобных средств и в целом терминологической базы, связанной с применением такой технологии в военных целях, позволит получить более ясное представление о предмете и перспективах обсуждения этой темы.

Рабочее определение должно отвечать следующим требованиям:

- а) содержать описание типов систем вооружений и военной техники с использованием ИИ, отдельные важные особенности их применения;
- б) не ограничиваться текущим пониманием указанных средств, а учитывать возможности их перспективного развития;
- в) быть универсальным с точки зрения понимания экспертным сообществом, включающим ученых, инженеров, техников, военнослужащих, юристов и специалистов по этике;

г) не толковаться как ограничивающее технический прогресс и наносящее ущерб исследованиям в области мирной робототехники и ИИ;

д) не определять системы вооружений и военную технику с использованием технологии ИИ исключительно через описание их функций.

Стоит избегать разделения обсуждаемых средств на «плохие» и «хорошие», то есть допускать их классификацию, исходя из политических предпочтений отдельной группы государств.

Действующие системы военного назначения высокой степени автоматизации не следует выделять в «особую» категорию, нуждающуюся в немедленных ограничениях и запретах. Именно такой уровень автоматизации позволяет подобным системам эффективно действовать в динамичной боевой обстановке и в различных средах, обеспечивая при этом надлежащую степень избирательности и точности и, как следствие, их соответствие принципам и нормам международного права, включая международное гуманитарное право (МГП).

Системы вооружений и военная техника с использованием технологии ИИ в контексте международного права

Общепризнано, что существующее международное право, в том числе МГП, полностью применяется к системам вооружений с использованием технологии ИИ.

Российская Федерация исходит из отсутствия на сегодняшний день убедительных оснований для введения каких-либо новых ограничений и запретов в отношении оружейных систем с применением технологии ИИ, модернизации или адаптации международного права, включая МГП, в связи с такими средствами. Преждевременны также дискуссии по согласованию неких «правил поведения» или норм и принципов «ответственного» применения в отношении систем вооружений и военной техники с использованием ИИ. Продвигаемая странами Запада концепция «ответственного» применения ИИ

строится на весьма спорных критериях, не известных международному праву (включая МГП), вызывает немало вопросов и не пользуется консенсусной поддержкой мирового сообщества.

Принципы гуманности, требования общественного сознания, равно как и правочеловеческий компонент не могут использоваться в качестве абсолютного и единственного достаточного условия для введения ограничительно-запретительных режимов в отношении конкретных видов вооружений и военной техники. Снятие озабоченностей в отношении систем вооружений и военной техники с использованием технологии ИИ лежит в плоскости добросовестного выполнения уже действующих международно-правовых норм.

Неукоснительное соблюдение норм и принципов международного права, включая МГП, в ситуациях вооружённых конфликтов остаётся одним из приоритетов Российской Федерации. Вооружённые Силы Российской Федерации строго руководствуются закреплёнными в федеральных и ведомственных правовых актах нормами МГП. В распорядительных документах и учебных программах для всех категорий военнослужащих отражаются вопросы соблюдения МГП, в том числе касающиеся применения новых видов вооружений. В 2022 году принята «Концепция деятельности Вооружённых Сил Российской Федерации в сфере разработки и применения систем вооружений с использованием технологий искусственного интеллекта».

Российское законодательство в полном объёме учитывает руководящие принципы в отношении систем вооружений с использованием технологии ИИ, одобренные консенсусом государствами-участниками КНО в 2019 г. Воспринимаем дальнейший обмен информацией о конкретных практических мерах по реализации указанных руководящих принципов на национальном уровне как способ укрепления доверия и повышения транспарентности.

Осуществление контроля над системами вооружения и военной техникой с применением технологии ИИ

В качестве важного ограничителя рассматриваем обеспечение человеческого контроля за функционированием систем вооружений и военной техники с применением ИИ. Для этого в контуре управления указанных средств должна быть предусмотрена возможность вмешательства человека-оператора или системы управления верхнего уровня для изменения режима функционирования таких систем, включая частичную или полную деактивацию.

Российская Федерация исходит из того, что человек всегда остается ответственным за решения о применении силы. Осуществляемый контроль основывается на всей имеющейся на момент принятия решения информации. При этом конкретные формы и методы человеческого контроля должны оставаться на усмотрение государств и не обязательно обеспечиваться исключительно путём прямого управления.

Контроль может осуществляться за счёт:

- а) повышения надёжности и отказоустойчивости системы;
- б) ограничения типов целей;
- в) ограничения по продолжительности функционирования, географическому охвату и масштабу применения;
- г) своевременного вмешательства и деактивации;
- д) проведения испытаний систем вооружений и военной техники с использованием технологии ИИ в реалистичных оперативных условиях;
- е) допуска к управлению (контролю) лиц, успешно освоивших процедуры использования средств с применением ИИ;
- ж) контроля процесса производства отдельных элементов и изделия в целом;
- з) контроля разбраковки и утилизации отдельных элементов и изделия в целом.

Считаем нецелесообразным введение в дискуссионный оборот продвигаемых отдельными государствами концепций «значимого

человеческого контроля», «формы и степени вовлечённости человека», «соответствующих контексту человеческих контроля и оценки», «предсказуемости, надёжности, отслеживаемости, объяснимости», поскольку подобные категории не имеют в целом отношения к праву и ведут исключительно к политизации дискуссий.

Ответственность

Российская Федерация исходит из того, что государства и отдельные лица (в том числе разработчики и производители) в любой момент времени несут ответственность в соответствии с международным правом за их решения по созданию и применению систем вооружений и военной техники с использованием технологии ИИ. Ответственность за применение таких средств ложится на должностное лицо, которое ставит им задачу и отдает приказ на их применение. При использовании систем вооружений и военной техники с задействованием ИИ такое лицо должно обладать надлежащими знаниями и навыками по их функционированию и эксплуатации, а также выполнять функции принятия решения о целесообразности применения, планирования форм и способов применения указанных средств.

Возможности и ограничения систем вооружений и военной техники с использованием технологии ИИ

Общеизвестно, что системы вооружений и военная техника с применением ИИ могут демонстрировать большую эффективность, чем человек-оператор при решении поставленных задач, сокращать вероятность упущений. В частности, подобные средства способны существенно снизить негативные последствия применения в контексте международного права, включая МГП, связанные с ошибками оператора, его психическим и физиологическим состоянием, нравственными, религиозными, моральными установками. Их применение может обеспечить повышение точности наведения средств поражения, направленных против военных объектов, способствовать сокращению вероятности

непреднамеренного нанесения ударов по гражданскому населению и гражданским объектам.

Оценка потенциальных рисков, связанных с применением систем вооружений и военной техники с использованием ИИ, и меры по их уменьшению должны быть частью цикла проектирования, разработки, испытания и развёртывания новых технологий в любых системах военного назначения.

Минимизация рисков в связи с подобными средствами могла бы осуществляться путём:

- а) эффективного управления жизненным циклом;
- б) организации разносторонних испытаний на всех этапах жизненного цикла, в том числе в условиях, близких к реальной обстановке;
- в) обеспечения надёжности и отказоустойчивости;
- г) определения критериев готовности;
- д) обеспечения максимальной защищённости от несанкционированного доступа;
- е) подготовки операторов;
- ж) приоритета в использовании технологии ИИ в процессе сбора и обработки информации, необходимой для поддержки принятия решений в военной сфере;
- з) непрерывного контроля за действиями таких систем со стороны оператора, обеспечение возможности экстренного прекращения выполнения боевой задачи по его команде;
- и) исключения попадания негосударственным субъектам, которые могут использовать их в противоправных целях.

Данные меры могут приниматься на всех этапах и стадиях жизненного цикла (разработка, производство, эксплуатация, утилизация) образцов вооружения, военной и специальной техники.

Дальнейшие действия

Считаем полезным продолжить рассмотрение между государствами вопросов, связанных с применением технологии ИИ в военных целях, в ГПЭ по САС как оптимальной международной площадке для такой дискуссии, в рамках действующих режимов в области КВРН, а также КОР. При этом соответствующее обсуждение на одной площадке не должно дублировать обмен мнениями, который уже ведётся в параллельных диалоговых форматах.

Выступаем против фрагментации усилий на данном направлении. Перенос проблематики применения технологии ИИ в военных целях на любые другие международные площадки, создание дополнительных форумов для её рассмотрения или её обсуждение в узком составе, без участия подавляющего большинства государств-членов ООН (в т.ч. из числа основных разработчиков систем вооружений с использованием ИИ, включая Российскую Федерацию) представляются контрпродуктивными.

В частности, деструктивный характер носит дискуссия по указанной тематике в рамках организуемых группой западных государств неинклюзивных «саммитов по ответственному применению ИИ в военных целях», а также саммитов по ИИ в целом. Данные мероприятия и принимаемые по их итогам документы не учитывают мнение всех заинтересованных сторон и не могут рассматриваться как основа для дальнейшей работы, которая бы отражала общее понимание данной тематики. Они имеют разобщающий эффект и не способствуют объединению усилий на указанном направлении.

Попытки «застолбить» односторонние подходы к указанной проблематике на альтернативных площадках, в т.ч. в рамках подобных «саммитов», в обход профильных многосторонних форматов, будут иметь крайне негативные последствия. Они способны серьезным образом подорвать ведущуюся конструктивную и инклюзивную работу по тематике «военного» ИИ и внести

раскол в усилия по выработке общих пониманий и рекомендаций на этом направлении.

В ходе дискуссии на указанных международных форумах полагаем необходимым сосредоточить основное внимание на согласовании единой профильной терминологической базы, а также подходов в отношении применения действующего международного права, включая МГП, к системам вооружений и военной технике с использованием технологии ИИ, обеспечения человеческого контроля за такими средствами и создаваемых данной технологией рисков и возможностей.

Российская Федерация просит Вас, господин Генеральный секретарь, учесть представленные выше предложения в Вашем содержательном докладе во исполнение пункта 7 резолюции ГА ООН 79/239, а также включить настоящий документ в приложение к Вашему докладу.