Позиция Российской Федерации в отношении перспектив рассмотрения проблематики военного применения искусственного интеллекта на международном уровне, в том числе в рамках ООН

Российская Федерация придает важное значение проблематике военного применения искусственного интеллекта (ИИ) и ее рассмотрению на международном уровне. Наши принципиальные подходы к данному вопросу изложены в национальном вкладе к докладу Генсекретаря ООН 80/78 о возможностях и проблемах в сфере обеспечения международного мира и безопасности в связи с использованием ИИ в военной сфере.

Параметры диалога по «военному» ИИ

В целом не возражаем против налаживания многостороннего диалога по вопросам применения ИИ в военной области и изучения последствий использования соответствующих средств для международного мира и безопасности, как это предлагается в указанном докладе Генсекретаря ООН (п. 45).

При этом исходим из того, что такой диалог должен строиться на основе консенсуса и инклюзивности, а также носить межгосударственный характер. Участие негосударственных субъектов должно осуществляться без ущерба для данного процесса, носить вспомогательный и деполитизированный характер, а также реализовываться исключительно в рамках компетенции указанных субъектов.

Эти фундаментальные принципы призваны гарантировать учет интересов каждого государства-члена, а также обеспечить дееспособность и эффективность принимаемых решений. Важность такого подхода подтверждается и в докладе Генсекретаря ООН, где, в частности, отмечаются необходимость выстраивания работы на основе консенсуса (п. 54), участие всех государств в обсуждении тематики «военного» ИИ (пп. 48, 63, 66), а также исключительная прерогатива государств в принятии решений (п. 48).

Площадка для диалога по ИИ в военной сфере

Рассматриваем разоруженческий механизм ООН в качестве эффективной и инклюзивной платформы. Он должен играть центральную роль в будущих дискуссиях по тематике применения ИИ в военной области на площадке Всемирной Организации, что отмечается в докладе Генсекретаря ООН (п. 52). Благодаря своему универсальному членскому составу, дискуссионному мандату и специализированному характеру деятельности Комиссия ООН по разоружению (КОР) идеально подходит для такого диалога. Солидный опыт КОР в разработке рекомендаций по различным аспектам контроля над вооружениями, разоружения и нераспространения (КВРН), а также наличие у Комиссии определенных наработок по вопросам ИИ в военной сфере позволили бы обеспечить конструктивную дискуссию и согласование практических руководящих принципов и рекомендаций по применению ИИ в военных целях.

Принципиально важно, чтобы вопросы «военного» ИИ, рассматриваемые на площадке Всемирной Организации, не дублировали темы, обсуждаемые в профильных форматах, прежде всего, в Группе правительственных экспертов (ГПЭ) государств-участников Конвенции о «негуманном» оружии (КНО) по смертоносным автономным системам вооружений (САС). Рассмотрение вопросов военного применения ИИ в рамках ГПЭ 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 данных технологий в военных целях.

Благодаря накопленному в рамках ГПЭ опыту взаимодействия, способности выдерживать разумный баланс между гуманитарными озабоченностями и законными оборонными интересами государств, а также участию ключевых стран-разработчиков Группа считается оптимальной площадкой для предметной работы с целью сближения позиций по вышеупомянутым аспектам «военного» ИИ (пп. 47, 50 доклада Генсекретаря

ООН). Важно продолжить данную работу в ГПЭ в рамках нового мандата после 2026 г.

Изучение упоминаемых в докладе Генсекретаря ООН (пп. 52, 54) опций относительно возможного рассмотрения тематики «военного» ИИ на Конференции по разоружению (КР) или в рамках специализированного процесса, в т.ч. новой профильной Рабочей группы открытого состава, считаем нецелесообразным. В случае с КР попытка «обогатить» ее повестку дня новой темой способна еще больше усугубить положение дел на форуме, где на протяжении длительного времени наблюдается отсутствие прогресса по программе работы. Инициирование специализированного процесса потребует согласования его организационных модальностей, а также выделения дополнительных финансовых ресурсов из бюджета ООН, что в условиях кризиса ликвидности во Всемирной Организации представляется нерациональным.

Одновременно отмечаем, что форматы с ограниченным составом участников, продвигающие односторонние подходы к «военному» ИИ, включая саммиты «ответственного применения» данных технологий (REAIM; упоминаются в пп. 41, 42 доклада Генсекретаря ООН), носят деструктивный характер. Они призваны подменить реальную дискуссию между государствами на профильных площадках и выработать в узком кругу некие понимания и стандарты в данной области, которые далее будут навязываться мировому сообществу в качестве готовых решений для одобрения. Региональные инициативы по ИИ в военной области (п. 43 доклада Генсекретаря ООН) охватывают ограниченное число участников, и их деятельность объективно не может носить универсальный характер. Продукты встреч в упомянутых форматах ограниченного состава не учитывают позиции всех заинтересованных стран и не могут рассматриваться как основа для дальнейшей работы. Кроме того, такие мероприятия чреваты фрагментацией усилий и могут подорвать инклюзивную работу по указанной проблематике на профильных площадках.

Полагаем необходимым в перспективе при обсуждении тематики «военного» ИИ обеспечить ее взаимодополняемость с диалоговыми форматами по КВРН, избегать создания произвольных механизмов международного надзора или дискриминационного экспортного контроля (п. 32 доклада Генсекретаря ООН).

Международно-правовое регулирование применения ИИ в военной сфере

На данный момент международное сообщество находится в самом начале пути обсуждения на профильных многосторонних площадках тематики применения новых технологий в военной сфере. Согласования требует понятийный аппарат, включая определение самого термина «ИИ». До выработки консенсусного «базового языка» преждевременно начинать дискуссии относительно ограничений и запретов в отношении соответствующих систем вооружений. Мнение о необходимости избегать преждевременного принятия мер регулирования зафиксировано в п. 46 доклада Генсекретаря ООН. Считаем также недопустимым введение ограничений, препятствующих законной инновационной деятельности и техническому прогрессу (п. 47 доклада Генсекретаря ООН).

Перспективным в рамках диалога по военным аспектам применения ИИ представляется анализ действующей международно-правовой базы в этой области. На сегодняшний день подавляющее большинство государств-членов согласны с тем, что нормы международного права в целом и международного гуманитарного права (МГП) в частности применимы к использованию ИИ в военной сфере, и в международном праве отсутствует категорический запрет на использование таких средств (п. 28 доклада Генсекретаря ООН). Действующее международное право, включая МГП, содержит широкий круг ограничений, позволяющих обеспечить гуманизацию боевых действий.

Государствам необходимо не только исходить из того, что военное использование ими ИИ должно соответствовать нормам международного права, но и иметь общее понимание, каким образом международное право

применяется к технологиям ИИ в военной сфере. Единое и четкое представление о практических особенностях осуществления действующих норм и принципов международного права в отношении «военного» ИИ является необходимым условием для выполнения государствами их обязательств.

При этом важно помнить, что технологии ИИ носят нейтральный характер, и диалог должен строиться не столько вокруг технологий как таковых, сколько вокруг аспектов, связанных с применением ИИ в военной сфере (п. 27 доклада Генсекретаря ООН). При этом соответствующее международно-правовое регулирование не должно препятствовать технологическому развитию или ограничивать равноправный доступ к ИИ для развивающихся стран.

Концепция «ответственного применения» «военного» ИИ, упомянутая в пп. 33-35 доклада Генсекретаря ООН, не может рассматриваться в качестве возможной основы для всеобъемлющей дискуссии по указанной тематике. Данная концепция основывается на не пользующихся консенсусной поддержкой принципах. Нормы и правила «ответственного применения» не являются общепризнанной категорией международного права. Они носят субъективный характер и не содержат четкого понимания, кем и на основании каких критериев будет определяться, является ли использование тем или иным государством таких технологий «ответственным» или «безответственным». Упомянутая концепция призвана закрепить лидирующие позиции определенной группы стран в данной сфере посредством навязывания мировому сообществу удобных исключительно таким государствам правил и стандартов, которые предоставляли бы полную «свободу рук» и сдерживали развитие соответствующих технологий в «неугодных» странах. Изложенную позицию разделяют и другие государства (п. 46 доклада Генсекретаря ООН).

Рассмотрение в рамках диалога этических аспектов и международного права прав человека (МППЧ; пп. 28 и 57 доклада Генсекретаря ООН) также представляется контрпродуктивным. Данные аспекты не могут

использоваться в качестве абсолютного и достаточного условия для введения каких-либо ограничительно-запретительных режимов в отношении указанных видов вооружений. Их внедрение в дискуссию может затруднить работу, связанную с прикладным анализом МГП, с учетом наличия споров в отношении его взаимодействия с МППЧ.

Считаем нецелесообразным введение в дискуссионный оборот таких концепций как «надежность, объяснимость, отслеживаемость в целях предотвращения непредвиденных последствий и сбоев» (п. 36 доклада Генсекретаря ООН). Подобные категории носят субъективный, неправовой характер и лишь осложняют усилия государств по выработке общих пониманий в отношении систем вооружений и военной техники с применением ИИ.

Преимущества систем вооружений и военной техники с применением ИИ

В докладе Генсекретаря ООН изложены преимущества технологий с использованием «военного» ИИ, в т.ч. способствующие лучшему соблюдению международного права, включая МГП.

Главным преимуществом ИИ называется быстродействие, в т.ч. при анализе информации и принятии решений. К достоинствам ИИ также относятся «надежность, безопасность и отказоустойчивость» (п. 7 доклада Генсекретаря ООН), а также повышение эффективности, точности и безошибочности операций (п. 56).

ИИ позволяет осуществлять эффективный анализ больших массивов данных, содействовать выявлению угроз, повышению степени ситуационной осведомленности и точности при проведении операций, а также способствовать уменьшению рисков для гражданского населения и усилению защиты гражданских объектов (п. 8). ИИ также снижает риски для личного состава, находящегося на передовой, поскольку средства с применением данных технологий могут заменить человека при ведении военных действий в труднодоступных или опасных районах (п. 12). ИИ оптимизирует процессы

управления и тылового обеспечения, а также профилактического обслуживания сложных систем (п. 11 доклада).

ИИ также способен улучшить координацию и связь (п. 9 доклада Генсекретаря ООН), укрепить безопасность информационно-коммуникационных технологий путем обнаружения вторжения или других злонамеренных действий, в т.ч. для защиты объектов критически важной инфраструктуры. Кроме того, данные технологии могут облегчать процесс выявления контента, используемого для распространения ложной информации и дезинформации (п. 10 доклада).

ИИ способствует обеспечению соблюдения норм МГП, в частности, его основополагающих принципов избирательности, пропорциональности и предосторожности при нападении, а также содействует защите гражданского населения и гражданских объектов. ИИ способен повышать степень ситуационной осведомленности в целом и понимание гражданской обстановки в частности, а также увеличивать точность и уменьшать риск человеческой ошибки (п.13 доклада).

Риски «военного» ИИ и их урегулирование

В целом не возражаем против того, чтобы в рамках международного диалога по применению ИИ в военной сфере наряду с анализом преимуществ соответствующего задействования данных технологий также были рассмотрены и связанные с этим риски, как предлагается в п. 46 доклада Генсекретаря ООН.

Максимальное снижение негативных последствий вооруженного конфликта путем ограничения его сторон в выборе средств и методов осуществления боевых операций является одной из задач международного права. При этом данные требования относятся к человеку, а не машине. Соответственно, одним из элементов минимизации возможных рисков, связанных с технологиями ИИ в военной сфере, является добросовестное соблюдение государствами международного права, включая МГП.

Среди ключевых положений МГП, призванных снизить риски, связанные с новыми видами оружия (к их числу можно отнести и технологии «военного» ИИ), особое значение имеет ст. 36 Дополнительного протокола I к Женевским конвенциям 1949 г., в соответствии с которой при изучении, разработке, приобретении или принятии на вооружение новых видов оружия, средств или методов ведения войны государство должно определить, подпадает ли их использование под запреты и ограничения, содержащиеся в применимых профильных международных договорах. Упомянутая статья не содержит положений о том, как именно такие обзоры должны проводиться и не налагает на страны обязательство информировать об их итогах, при этом не исключает возможность добровольного обмена наилучшими практиками их выполнения. Это позволяет обеспечить гибкий и эффективный подход, учитывающий национальную специфику и позиции государств, их потенциал и ограничения в технологической сфере.

На решение возможных вызовов, связанных с задействованием технологий «военного» ИИ, направлено обеспечение человеческого контроля за функционированием технологий с использованием ИИ в военной сфере. Его конкретные формы закладываются государствами на практике на национальном уровне и не обязательно могут применяться исключительно путем прямого управления. Среди возможных способов обеспечения человеческого контроля – повышение надежности и отказоустойчивости системы; ограничение типов целей; ограничения по продолжительности функционирования, географическому охвату и масштабу применения; своевременное вмешательство и деактивация; проведение испытаний в реалистичных оперативных условиях; допуск к управлению (контролю) лиц, успешно освоивших процедуры применения соответствующих систем; контроль процесса производства отдельных элементов и изделия в целом; контроль разбраковки и утилизации отдельных элементов и изделия в целом.

Всеобъемлющий обмен мнениями на профильных площадках, прежде всего, в рамках ГПЭ по САС, также призван способствовать минимизации

вызовов в связи с «военным» ИИ, в т.ч. через согласование конкретных рекомендаций и руководящих принципов. В ходе сессий данной Группы государствами неоднократно на конкретных примерах подчеркивалась роль технологий ИИ в военной сфере при решении задач обороны, борьбы с террористической угрозой, минной опасностью и т.д.

Указанные способы обеспечения человеческого контроля способны эффективно снизить риски, включая те, которые упомянуты в докладе Генсекретаря ООН (пп. 16, 18, 19). Представленные в докладе способы снижения рисков в связи с технологиями «военного» ИИ (п. 37) также представляются релевантными.

Темы для дальнейшего обсуждения

Помимо уже зафиксированных в настоящем рабочем документе предложений по анализу применения действующей международно-правовой базы в отношении «военного» ИИ, рисков и преимуществ, ряд изложенных в п. 49 доклада Генсекретаря ООН тем также могли бы представлять интерес для дальнейшего диалога. В частности, речь идет о дискуссии по обеспечению соблюдения норм международного права, в т.ч. МГП; выработке общего понимания определений и терминологии; мерам обеспечения транспарентности и укреплению доверия; вопросу об автономии в сфере применения силы; системам ИИ, непосредственно поддерживающим боевые операции; обеспечению адекватных механизмов управления данными; укреплению международного сотрудничества и помощи; поддержке наращивания потенциала, в т.ч. путем обмена знаниями, передачи технологий и обмена передовым опытом, с целью преодоления отставаний в сфере цифровых технологий и ИИ. Требуется также изучить вопрос четкого разграничения «гражданского» и «военного» ИИ, как это предлагается в п. 61 доклада Генсекретаря ООН.

При возможном обмене мнениями по вопросам международного взаимодействия особое внимание необходимо уделить тематике обеспечения недискриминационного и справедливого доступа всех государств, особенно

стран Глобального Юга, к технологиям ИИ. Доступ к технологиям для целей развития не должен ограничиваться. Возможный обмен информацией и опытом в контексте предлагаемого обсуждения мер транспарентности в отношении военного применения ИИ должен осуществляться государствами сугубо на добровольной основе.

Предложение по обеспечению на международном уровне «содействия национальному регулированию, в т.ч. ради обеспечения соблюдения норм международного права частным сектором» (п. 49 доклада Генсекретаря ООН) является контрпродуктивным. Внутренняя регламентация прикладных вопросов, связанных с соблюдением международного права в сфере «военного» ИИ, в т.ч. частным сектором, должна оставаться на усмотрение государств. Обсуждение упомянутых в п. 49 доклада Генсекретаря ООН аспектов «защиты человеческого достоинства и прав человека» и «содействия продолжению регионального диалога» представляется нецелесообразным с учетом пониманий, представленных в предыдущих параграфах рабочего документа.

Крайне важно, чтобы формулируемый возможный перечень перспективных тем для диалога осуществлялся на базе документов, поддерживаемых абсолютно всеми государствами. Так, ряд стран дистанцировался либо проголосовал против упоминаемых в докладе Генсекретаря ООН «Пакта будущего» (резолюция 79/1 ГА ООН; п. 38), резолюций Генассамблеи 78/241, 79/62 «САС» (п. 39), а также резолюции 79/239 «Применение ИИ в военной области и его последствия для международного мира и безопасности» (п. 38). Их продвижение в качестве основы для международного взаимодействия по тематике ИИ лишь осложнит конструктивную работу.

Изложенные соображения призваны содействовать дальнейшему эффективному содержательному рассмотрению проблематики военного применения ИИ на международном уровне, в т.ч. в рамках ООН. Рассчитываем на всесторонний учет предложений Российской Федерации, в

т.ч. при подготовке резюме дискуссий по итогам неформальных консультаций, организованных во исполнение резолюции ГА ООН 80/58.