



# Генеральная Ассамблея

Distr.: General

5 June 2025

Russian

Original: Chinese/English/French/  
Russian/Spanish

Восьмидесятая сессия

Пункт 101 первоначального перечня\*

Всеобщее и полное разоружение

## Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности

### Доклад Генерального секретаря

#### *Резюме*

В настоящем докладе содержится резюме элементов мнений, полученных от государств-членов и государств-наблюдателей в соответствии с резолюцией [79/239](#), составленное без ущерба для их индивидуальных позиций. В нем описываются возможности и проблемы, связанные с применением искусственного интеллекта в военной области, приводится каталог существующих и новых нормативных предложений, излагается обзор инициатив по вопросу о применении искусственного интеллекта в военной области, описываются соображения касательно дальнейших шагов, а также приводятся замечания и выводы Генерального секретаря.

\* [A/80/50](#).



## Содержание

	<i>Стр.</i>
I. Введение . . . . .	4
II. Справочная информация . . . . .	4
III. Возможности и проблемы . . . . .	5
IV. Существующие и новые нормативные предложения. . . . .	8
V. Инициативы по вопросу о применении искусственного интеллекта в военной области	11
VI. Дальнейшие шаги . . . . .	12
VII. Замечания и выводы Генерального секретаря . . . . .	14
Приложение I	
Полученные ответы. . . . .	17
A. Государства-члены . . . . .	17
Аргентина. . . . .	17
Австрия. . . . .	18
Чили. . . . .	22
Китай. . . . .	25
Египет . . . . .	27
Сальвадор . . . . .	30
Финляндия . . . . .	32
Франция . . . . .	35
Германия. . . . .	38
Греция . . . . .	42
Индия . . . . .	45
Индонезия. . . . .	46
Иран (Исламская Республика) . . . . .	50
Израиль. . . . .	51
Италия. . . . .	53
Япония . . . . .	56
Литва. . . . .	59
Мексика. . . . .	61
Нидерланды (Королевство) . . . . .	64
Новая Зеландия . . . . .	69
Норвегия. . . . .	70
Пакистан. . . . .	75
Перу. . . . .	80
Республика Корея . . . . .	83

Российская Федерация . . . . .	88
Сербия . . . . .	93
Сингапур . . . . .	95
Испания . . . . .	97
Швейцария . . . . .	101
Украина . . . . .	106
Соединенное Королевство Великобритании и Северной Ирландии . . . . .	108
B. Европейский союз . . . . .	113

## Приложение II

Replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry . . . . .	115
A. International and regional organizations . . . . .	115
African Commission on Human and Peoples' Rights . . . . .	115
B. International Committee of the Red Cross . . . . .	119
C. Civil society . . . . .	124
Autonorms . . . . .	124
Global Commission on Responsible Artificial Intelligence in the Military Domain . . . . .	127
InterAgency Institute . . . . .	132
International Committee for Robot Arms Control . . . . .	134
International Humanitarian Law and Youth Initiative . . . . .	135
Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand . . . . .	139
Ploughshares . . . . .	142
Soka Gakkai International . . . . .	144
Stop Killer Robots . . . . .	145
Stop Killer Robots Youth Network . . . . .	148
Unione degli Scienziati Per Il Disarmo . . . . .	152
Women's International League for Peace and Freedom . . . . .	154
D. Scientific community . . . . .	155
AI, Automated Systems, and Resort-to-Force Decision Making Research Project, the Australian National University . . . . .	155
Queen Mary University of London, T.M.C. Asser Institute, University of Southern Denmark and University of Utrecht . . . . .	161
United Nations Institute for Disarmament Research . . . . .	165
E. Industry . . . . .	169
Microsoft . . . . .	169

## I. Введение

1. В пункте 7 резолюции [79/239](#) о применении искусственного интеллекта (ИИ) в военной области и его последствиях для международного мира и безопасности Генеральная Ассамблея просила Генерального секретаря запросить мнения государств-членов и государств-наблюдателей о возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением ИИ в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений, и представить Генеральной Ассамблее на ее восьмидесятой сессии для дальнейшего обсуждения государствами субстантивный доклад, содержащий обобщение этих мнений и каталог существующих и новых нормативных предложений, а также приложение с изложением этих мнений. В пункте 8 той же резолюции Ассамблея просила также Генерального секретаря запросить мнения международных и региональных организаций, Международного комитета Красного Креста, гражданского общества, научного сообщества и отраслевых кругов и включить эти мнения на том языке, на котором они были представлены, в приложение к вышеупомянутому докладу. Настоящий доклад представляется во исполнение этих просьб.

2. 12 февраля 2025 года Управление по вопросам разоружения направило всем государствам-членам и государствам-наблюдателям вербальную ноту, в которой обратило их внимание на пункт 7 резолюции [79/239](#) Генеральной Ассамблеи и предложило представить мнения по этому вопросу. Структурам, указанным в пункте 8 этой резолюции, также были направлены вербальные ноты и письма, в которых их внимание было обращено на этот пункт и было предложено представить мнения по этому вопросу. Мнения, полученные к 11 апреля 2025 года, приводятся в приложениях к настоящему докладу. Любые мнения, полученные после этого крайнего срока, будут опубликованы на веб-сайте Управления по вопросам разоружения на том языке, на котором они были представлены.

3. В разделах II–VI настоящего доклада приводится резюме элементов мнений, полученных от государств-членов и государств-наблюдателей, составленное без ущерба для их индивидуальных позиций. Замечания и выводы Генерального секретаря изложены в разделе VII.

## II. Справочная информация

4. Государства указали на стремительное развитие науки и техники в целом и ИИ в частности, отметив их широкое воздействие на общество. В частности, государства отметили, что ИИ способен изменить все аспекты военного дела и оказать значительное воздействие на международный мир и безопасность.

5. Несколько государств сообщили о том, какие приложения ИИ применяются в военной области в настоящее время, а также о своих собственных усилиях по применению ИИ в военных операциях. Признавая важность обсуждения вопросов, касающихся смертоносных автономных систем вооружений, государства отметили, что вопрос о применении ИИ в военной области носит более широкий характер и охватывает более широкий спектр возможностей.

### III. Возможности и проблемы

6. Было отмечено, что искусственный интеллект открывает возможности и создает проблемы, к которым следует подходить с реалистических позиций. Было признано, что в настоящее время с учетом темпов развития ИИ все возможности и проблемы предвидеть невозможно. Было высказано мнение о том, что саму эту технологию порицать не следует.

#### A. Возможности

7. Было отмечено, что главным преимуществом ИИ является быстроедействие, в том числе при анализе информации и принятии решений. В качестве еще одного преимущества была отмечена масштабность, поскольку ИИ может выступать в качестве фактора повышения эффективности. Несколько государств отметили, что ИИ может повышать степень эффективности, точности и безошибочности операций, что позволит уменьшить вероятность ошибок по сравнению с людьми. В числе дополнительных характеристик были отмечены надежность, безопасность и отказоустойчивость.

##### *Виды применения*

8. Несколько государств отметили, что ИИ применяется в целях ведения разведки, наблюдения и рекогносцировки, где его можно использовать для эффективного анализа больших массивов данных, содействия выявлению угроз, повышения степени ситуационной осведомленности и повышения точности в проведении операций. Было отмечено, что эти же особенности позволяют ИИ поддерживать процесс принятия решений и систему командования и управления, что потенциально может способствовать повышению точности в проведении операций, уменьшению рисков для гражданского населения и усилению защиты гражданских объектов. Вместе с тем было также подчеркнуто, что средства ИИ не должны заменять собой процесс принятия решений человеком.

9. Несколько государств указали, что ИИ можно внедрять в безоператорные системы. Было отмечено, что ИИ способен улучшить координацию и связь между военными структурами, а также между военными и иными структурами, например, поставщиками гуманитарной помощи. В целом было отмечено, что ИИ может облегчить бремя выполнения рутинных или повторяющихся операций и расширить возможности человека при решении сложных задач.

10. По мнению некоторых государств, ИИ можно использовать для усиления безопасности информационно-коммуникационных технологий путем обнаружения вторжений или других злонамеренных действий, в том числе для защиты объектов критически важной инфраструктуры. Было отмечено, что ИИ может использоваться для выявления созданного ИИ контента, используемого для распространения ложной информации и дезинформации, а также для обнаружения ненавистнических высказываний, пропаганды или изменений в общественных настроениях.

11. Были упомянуты и другие сферы применения ИИ, не имеющие непосредственного отношения к боевым действиям, включая оптимизацию тылового обеспечения, профилактического обслуживания, закупок, распределения ресурсов, управления, моделирования и обучения.

*Вопросы международного мира и безопасности*

12. Несколько государств сочли, что ИИ может способствовать поддержанию международного мира и безопасности, например, обеспечение ситуационной осведомленности при поддержке ИИ может помочь в уменьшении рисков и способствовать деэскалации конфликтов. Было отмечено, что применение ИИ может уменьшать риск для личного состава, например, заменяя людей в процессе решения некоторых опасных задач, таких как утилизация неразорвавшихся боеприпасов, или поддерживая поисково-спасательные операции в удаленных районах.

13. Было высказано мнение о том, что ИИ может способствовать обеспечению соблюдения норм международного гуманитарного права, в частности, его основополагающих принципов избирательности, пропорциональности и предосторожности при нападении, а также обеспечению защиты гражданского населения и гражданских объектов. В этой связи несколько государств отметили способность ИИ повышать степень ситуационной осведомленности в целом и понимание гражданской обстановки в частности, а также повышать точность и уменьшать риск человеческой ошибки. Кроме того, было отмечено, что ИИ может способствовать расследованию случаев гибели гражданских лиц и тем самым обеспечивать привлечение виновных к ответственности.

14. Несколько государств высказали мнение о том, что ИИ может способствовать мониторингу и проверке осуществления соглашений о разоружении, нераспространении и контроле над вооружениями. Отмечался потенциал ИИ в плане поддержки миссий по поддержанию мира, в том числе в сфере планирования, материально-технического снабжения и наблюдения за прекращением огня. В числе других сопутствующих сфер применения ИИ можно назвать охрану границ, борьбу с терроризмом, выявление незаконных оружейных программ, а также оптимизацию гуманитарной помощи и реагирования на бедствия.

## **В. Проблемы**

15. Несколько государств отметили, что стремительное развитие новейших технологий в целом и искусственного интеллекта в частности порождает проблемы для международного мира и безопасности. Эти проблемы важно понимать, однако в настоящее время невозможно полностью предвидеть их в полном объеме.

16. Были отмечены следующие проблемы, порождаемые ИИ:

- ускорение цикла наблюдения, оценки обстановки, принятия решения и нанесения удара с соответствующим сокращением времени на принятие решений;
- повышение степени автономности и утрата человеческого контроля, особенно в контексте применения силы;
- потенциал неправомерного или злонамеренного применения;
- чрезмерное доверие людей к приложениям ИИ;
- углубление технологического неравенства государств.

*Вопросы международного мира и безопасности*

17. Несколько государств отметили, что внедрение ИИ в военной области может породить проблемы для международного мира и безопасности. Применение ИИ может привести к усилению риска недоразумений, просчетов и

непреднамеренной эскалации, в том числе в результате увеличения скорости и расширения масштабов осуществления операций при поддержке ИИ или в результате технических сбоев. Эти причины также могут привести к понижению порога применения силы. Несколько государств выразили обеспокоенность по поводу развязывания гонки вооружений в этой области. Было высказано предположение о том, что применение ИИ может сместить баланс от оборонительных действий к наступательным и что нарастание диспропорций между государствами может вызвать усиление нестабильности и тем самым подорвать международный мир и безопасность.

18. Несколько государств выразили обеспокоенность по поводу потенциально дестабилизирующего воздействия распространения средств ИИ, в том числе среди негосударственных субъектов. Было отмечено, что в настоящее время не существует многосторонней рамочной системы контроля за распространением вооружений с поддержкой ИИ.

#### *Технологические аспекты*

19. Государства рассмотрели риски, обусловленные технологическими аспектами, включая:

- технические отказы и неисправности;
- конструктивные недостатки;
- непреднамеренное поведение, отклоняющееся от проектных параметров;
- уязвимость к кибератакам и отравлению данных;
- предвзятость алгоритмов и данных, включая гендерную предвзятость;
- принятие необъективных решений в результате использования автоматизации недостаточно подготовленными операторами;
- проблемы конфиденциальности, возникающие в процессе сбора и обработки больших объемов персональных данных для обучения моделей ИИ;
- проблемы, вызванные низким качеством обучения моделей ИИ;
- проблемы, порождаемые некачественными процедурами тестирования, оценки, сертификации и проверки;
- ошибки выбора целей;
- чрезмерное энергопотребление;
- чрезмерная зависимость от внешних поставщиков.

20. Несколько государств выразили обеспокоенность по поводу прозрачности и объяснимости сложных систем ИИ, которые часто называют «черными ящиками». Кроме того, была выражена обеспокоенность по поводу использования таких гражданских приложений ИИ, как генеративный ИИ, которые могут привносить дополнительный уровень сложности и неопределенности в конфликтную ситуацию. Несколько государств выразили обеспокоенность и по поводу слияния ИИ и других технологий.

#### *Правовые и гуманитарные аспекты*

21. Несколько государств отметили, что ИИ создает проблемы в сфере соблюдения норм международного права, в частности международного гуманитарного права и международного права прав человека. Применение ИИ может привести

к неизбежному применению силы и поставить вопрос об ответственности и подотчетности в случае совершения незаконных или противоправных действий. Были затронуты такие сопутствующие вопросы, как защита гражданского населения и гражданской инфраструктуры, а также потенциал повышения интенсивности и смертоносности конфликтов для комбатантов.

22. Несколько государств затронули проблемы этического характера, отметив, что применение ИИ потенциально снижает уровень сострадания, нравственной аргументации и здравого смысла.

#### *Потенциальные сферы неправомерного применения*

23. Несколько государств отметили, что существует возможность применения ИИ для совершения кибератак как государственными, так и негосударственными субъектами, в том числе против объектов критически важной инфраструктуры. Кроме того, существует вероятность применения ИИ для проведения кампаний по распространению ложной информации и дезинформации, в том числе для генерирования дезинформации и дипфейков, а также для распространения ботов, управляемых ИИ. Использование такой ложной информации и дезинформации, например, для оказания влияния на выборы, может вызвать дестабилизацию.

#### *Оружие массового уничтожения*

24. Несколько государств подчеркнули важность сохранения человеческого контроля над ядерным оружием и системами его доставки и выразили обеспокоенность по поводу возможного внедрения ИИ в системы ядерного командования, управления и связи. Было отмечено, что некоторые государства, обладающие ядерным оружием, взяли на себя сохранять контроль и участие человека во всех действиях, имеющих критически важное значение для информационного обеспечения и осуществления суверенных решений, касающихся применения ядерного оружия. Были отмечены потенциальные последствия с точки зрения стратегической стабильности и эскалации.

25. Несколько государств выразили опасения по поводу того, что ИИ может способствовать распространению оружия массового уничтожения, в том числе среди негосударственных субъектов. В этом контексте особое беспокойство вызывает возможность использования ИИ для разработки и производства биологического оружия. Было подчеркнуто, что, согласно положениям действующих договоров, ИИ не должен использоваться в этих целях. Было также высказано мнение о том, что существует возможность использования ИИ для сдерживания распространения оружия массового уничтожения.

## **IV. Существующие и новые нормативные предложения**

26. Несколько государств заявили, что ИИ должен использоваться в мирных целях, в том числе для мирного урегулирования споров. Кроме того, государства подчеркнули важность изучения и уменьшения рисков, возникающих в результате применения ИИ в военной области, причем некоторые из них отметили, что проблемы, возникающие в результате применения военного ИИ, должны решаться совместными усилиями.

27. Государства призвали использовать при рассмотрении вопроса о применении ИИ в военной области подход, предусматривающий:

- гибкость, сбалансированность, реалистичность и поэтапность, и таким образом способность адаптироваться к технологическому прогрессу;



- предосторожность;
- ориентированность на весь жизненный цикл средств ИИ, включая этап подготовки проектного задания, проектирование, разработку, оценку, тестирование, внедрение, использование, продажу, закупку, эксплуатацию и вывод из эксплуатации;
- упор на применение и использовании ИИ, а не на саму технологию;
- учет существующих обязательств.

Было предложено в ходе работы над этим вопросом четко разграничивать смертоносные и несмертоносные виды применения.

#### *Правовые аспекты*

28. Государства сослались на резолюцию [79/239](#), в которой Генеральная Ассамблея подтвердила, что нормы международного права, включая Устав Организации Объединенных Наций, международное гуманитарное право и международное право прав человека, применяются к регулируемым ими вопросам, которые возникают на всех этапах жизненного цикла средств ИИ, в том числе к системам, функционирование которых в военной области обеспечивает ИИ. Было отмечено, что нормы международного права в целом и международного гуманитарного права в частности не содержат категорического запрета на использование средств ИИ.

29. Государства подтвердили, что они соблюдают нормы международного права в процессе применения ИИ в военной области. Было высказано мнение о том, что соблюдение правовых обязательств, особенно обязательств, вытекающих из норм международного права, должно быть ключевым аспектом в регулировании, проектировании и внедрении систем ИИ в военной области. Кроме того, было высказано мнение о том, что системы ИИ должны проектироваться в расчете на обеспечение более строгого соблюдения норм международного гуманитарного права. Несколько государств подчеркнули в этой связи важность проведения правовых обзоров новых вооружений, средств и методов ведения войны.

30. Некоторые государства подчеркнули важность учета не только законодательной базы, но и этических аспектов.

#### *Аспекты, касающиеся международного мира и безопасности*

31. Несколько государств отметили, что применение ИИ в военной области должно укреплять международный мир и безопасность и не должно вызывать нестабильность или эскалацию. Было высказано мнение о том, что государствам не следует стремиться к получению абсолютного военного преимущества с помощью ИИ и нужно обеспечивать, чтобы такая технология не стала инструментом вторжений и достижения гегемонии.

32. Несколько государств отметили, что ИИ не должен подрывать существующие соглашения о разоружении, нераспространении и контроле над вооружениями. Прозвучали призывы прилагать усилия по предотвращению распространения технологий ИИ среди негосударственных субъектов. Было подчеркнуто, что важно избегать создания произвольных механизмов международного надзора или дискриминационного экспортного контроля.

*Ответственное применение искусственного интеллекта в военной области*

33. Несколько государств высказали мнение о том, что ИИ следует применять ответственно на протяжении всего его жизненного цикла. Было отмечено, что концепция ответственности должна быть увязана с вопросами законности и подотчетности.

34. Несколько государств подчеркнули важность принятия человекоцентричного подхода к ИИ. Многие государства подчеркнули важность обеспечения постоянного контроля и ответственности со стороны человека. Была отмечена важность таких понятий, как «контекстуальный контроль со стороны человека и его участие в принятии решений» и «значимый контроль со стороны человека». Другие же государства отметили, что эти понятия недостаточно четко определены. Было высказано мнение о том, что использование понятия «значимого контроля со стороны человека» может помешать проведению законных исследований или неоправданно ограничить применение ИИ в военной области.

35. Государства подчеркнули важность обеспечения ответственности и подотчетности человека в соответствии с нормами международного права, в том числе в рамках ответственной системы командования и управления людьми.

*Технологические аспекты*

36. Государства рассмотрели технические аспекты принципов регулирования, в частности:

- защищенность в целях обеспечения устойчивости систем ИИ к внешним угрозам;
- безопасность, в том числе за счет встроенных ограничений для сведения вреда к минимуму;
- надежность в целях предотвращения непредвиденных последствий и сбоев;
- четкие операционные границы и ограничения в целях предотвращения непредвиденного поведения;
- четко проработанные сценарии применения;
- обеспечение управляемости для надлежащего взаимодействия человека с машиной и уменьшения предвзятости;
- равенство и справедливость;
- защита конфиденциальности;
- объяснимость, понятность и отслеживаемость;
- транспарентность.

37. Особое внимание было уделено вопросу об использовании учебных данных, обеспечивающих возможность полного соблюдения норм международного права. Несколько государств подчеркнули важность тестирования на протяжении всего жизненного цикла для выявления ошибок и обеспечения надежности. Было также подчеркнуто, что необходимо обеспечить надлежащую подготовку личного состава, работающего с ИИ, в целях уменьшения рисков и обеспечения соблюдения норм международного гуманитарного права. Была подчеркнута важность мониторинга работы систем на протяжении всего их жизненного цикла, а также принятия мер для надежного отключения систем на этапе их вывода из эксплуатации.

## V. Инициативы по вопросу о применении искусственного интеллекта в военной области

### *Международные форумы*

38. Некоторые государства отметили текущие обсуждения в Организации Объединенных Наций, а также Пакт во имя будущего (резолюция 79/1 Генеральной Ассамблеи) и содержащийся в приложении к нему Глобальный цифровой договор, а также резолюцию Генеральной Ассамблеи о применении искусственного интеллекта в военной области и его последствиях для международного мира и безопасности (резолюция 79/239). Было отмечено также заседание Совета Безопасности по формуле Аррии, посвященное вопросам использования безопасного, инклюзивного и заслуживающего доверия ИИ для поддержания международного мира и безопасности, которое состоялось 4 апреля 2025 года.

39. Государства отметили многосторонние обсуждения тем, связанных с применением ИИ в военной области, в частности обсуждения в Комиссии по разоружению в рамках пункта повестки дня, озаглавленного «Рекомендации относительно выработки общего понимания в вопросе о новейших технологиях в контексте международной безопасности», работу Группы правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений, созданной в соответствии с Конвенцией о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие, а также резолюции Генеральной Ассамблеи по смертоносным автономным системам вооружений (резолюции 78/241 и 79/62).

40. Государства упомянули также о своем участии в мероприятиях по вопросам применения ИИ в военной области, организованных Управлением по вопросам разоружения и Институтом Организации Объединенных Наций по исследованию проблем разоружения.

### *Инициативы, осуществляемые под руководством государств*

41. Несколько государств отметили, что они выступили в качестве организаторов или участников инициатив по вопросам применения ИИ в военной области, включая следующие:

- Политическая декларация об ответственном применении искусственного интеллекта и автономных систем в военных целях и процесс ее последующей реализации;
- процесс «Ответственное применение искусственного интеллекта в военной области», в рамках которого в 2023 году состоялись конференция в Королевстве Нидерландов, на которой был принят Призыв к действию, и конференция в Республике Корея в 2024 году, на которой был принят План действий. Глобальная комиссия по вопросу об ответственном применении искусственного интеллекта в военной области должна опубликовать доклад до следующей конференции, которая пройдет в Испании в 2025 году;
- саммит по искусственному интеллекту, состоявшийся во Франции в 2025 году, на котором была принята Парижская декларация о сохранении человеческого контроля в оружейных системах, использующих искусственный интеллект;
- Саммит по вопросам безопасности ИИ, который состоялся в Соединенном Королевстве Великобритании и Северной Ирландии в 2023 году и на котором была принята Блетчлийская декларация;

- работа по вопросам ИИ, проводимая в рамках Группы семи;
- Партнерство в сфере ИИ в интересах обороны;
- Глобальная инициатива по регулированию ИИ, выдвинутая в 2023 году.

42. Было высказано мнение о том, что эти инициативы являются полезными, однако они чреваты фрагментацией усилий. Кроме того, была выражена обеспокоенность тем, что результаты этих инициатив не учитывают мнения всех заинтересованных государств и могут подорвать инклюзивную работу в этой сфере.

#### *Региональные инициативы*

43. Государства отметили важность региональных инициатив по содействию проведению всесторонних обсуждений вопросов применения ИИ в военной области с учетом конкретных условий. В этом отношении были приведены следующие примеры:

- Совместное заявление о сотрудничестве в сфере, касающейся применения ИИ в оборонном секторе, принятое на выездном совещании министров обороны стран АСЕАН в 2025 году;
- шестнадцатая Конференция министров обороны стран Северной и Южной Америки 2024 года, на которой была принята Мендосская декларация;
- мероприятия в рамках Организации Североатлантического договора, включая ее стратегию по ИИ, последний раз пересмотренную в 2024 году, и ее принципы ответственного применения, разработанные в 2021 году;
- региональные консультации в контексте процесса «Ответственный искусственный интеллект в военной области» 2024 года, проведенные в Кении, Нидерландах (Королевство), Сингапуре, Турции и Чили.

#### *Национальные инициативы*

44. Государства рассказали о своих национальных усилиях, в том числе о действующем законодательстве, нормативных актах, стратегиях и органах по вопросам ИИ, а также об усилиях по их разработке.

## **VI. Дальнейшие шаги**

45. Государства призвали наладить диалог по вопросам применения ИИ в военной области. Несколько государств призвали продолжить изучение последствий применения ИИ в военной области на международный мир и безопасность.

46. Многие государства отметили, что целью дальнейшего диалога должно стать уменьшение рисков, порождаемых ИИ в военной области. Было высказано мнение о том, что цель диалога должна заключаться в разработке нормативно-правовой базы или системы регулирования. Несколько государств призвали заняться разработкой норм, правил и принципов, регулирующих жизненный цикл средств ИИ в военной области. Некоторые государства высказались за разработку юридически обязывающих рамок, однако другие сочли, что на данном этапе необходимости в принятии новых правовых мер не существует. Было также высказано мнение о том, что нормы, правила и принципы могли бы стать основой для принятия правовых обязательств на более позднем этапе. Несколько государств высказались против идеи принятия норм, правил и принципов ответственной разработки, внедрения или применения, отметив, что она не

получила консенсусной поддержки. Было высказано мнение о том, что следует избегать преждевременного принятия мер регулирования.

47. Была подчеркнута важность не допускать дублирования и фрагментации в сфере регулирования. Было высказано мнение о том, что при обсуждении вопросов регулирования следует сбалансированно учитывать гуманитарные соображения, соображения безопасности и развития. Государства подчеркнули, что важно избегать введения ограничений, которые препятствовали бы законной инноваторской деятельности и техническому прогрессу. По мнению нескольких государств, не следует препятствовать использованию ИИ в мирных целях, особенно развивающимися странами.

48. Было высказано мнение о том, что в рамках любого подхода к регулированию следует учитывать, что между государствами существуют различия в степени внедрения ИИ в военный потенциал и в ситуациях в плане безопасности. Была подчеркнута важность участия всех государств в обсуждении вопросов регулирования ИИ в военной области. Многие государства отметили, что будущие обсуждения должны проходить с участием многих заинтересованных сторон, включая международные и региональные организации, гражданское общество, научное сообщество и отраслевые круги. Вместе с тем было подчеркнуто, что принятие решений должно оставаться исключительной прерогативой государств.

49. Государства рассмотрели различные первоочередные задачи для будущего диалога по вопросам применения ИИ в военной области, включая:

- обеспечение соблюдения норм международного права, в частности международного гуманитарного права;
- защиту человеческого достоинства и прав человека;
- стремление к выработке общего понимания определений и терминологии;
- рассмотрение мер по обеспечению транспарентности и укреплению доверия;
- рассмотрение вопроса об автономии в сфере применения силы;
- рассмотрение вопроса о системах ИИ, непосредственно поддерживающих боевые операции;
- обеспечение адекватных механизмов управления данными;
- укрепление международного сотрудничества и помощи;
- поддержка наращивания потенциала, в том числе путем обмена знаниями, передачи технологий и обмена передовым опытом, с целью преодоления отставаний в сфере цифровых технологий и ИИ;
- содействие продолжению регионального диалога;
- содействие национальному регулированию, в том числе ради обеспечения соблюдения норм международного права частным сектором.

50. Несколько государств высказали мнение о том, что рассмотрение вопроса о смертоносных автономных системах вооружений должно стать частью любого обсуждения вопросов применения ИИ в военной области. Кроме того, было высказано мнение о том, что ведущиеся в настоящее время обсуждения вопросов, касающихся таких систем, дополняют дискуссии относительно применения ИИ в военной области. Несколько государств напомнили о своих позициях в

отношении смертоносных автономных систем вооружений<sup>1</sup>. Одни государства выразили мнение о том, что наиболее подходящим форумом для обсуждения вопросов, касающихся применения ИИ в военной области, является Группа правительственных экспертов, созданная в соответствии с Конвенцией по конкретным видам обычного оружия, а другие государства заявили, что эта группа не является подходящим форумом для таких обсуждений поскольку она обладает конкретным мандатом и не имеет универсального членского состава.

51. Несколько государств призвали обсуждать вопросы применения ИИ в военной области на форумах Организации Объединенных Наций. Было высказано мнение о том, что настоящий доклад мог бы послужить основой для таких обсуждений. Государства отметили, что будущие обсуждения должны дополнять такие текущие процессы, такие как деятельность в рамках Рабочей группы открытого состава по вопросам безопасности в сфере использования информационно-коммуникационных технологий и самих информационно-коммуникационных технологий.

52. Несколько государств высказали мнение о том, что разоруженческий механизм Организации Объединенных Наций представляет собой эффективную и инклюзивную платформу и должен играть центральную роль в будущих обсуждениях вопросов применения ИИ в военной области. Было предложено, чтобы Конференция по разоружению обсудила вопрос об ИИ, в частности применительно к ядерному оружию. Было высказано мнение о том, что обсуждения можно проводить и в Первом комитете Генеральной Ассамблеи, который мог бы поручить Генеральному секретарю регулярно представлять доклады о ходе технологического прогресса в сфере применения ИИ в военной области. Несколько государств предложили провести обсуждения в контексте Комиссии по разоружению.

53. Было также предложено провести обсуждения в контексте Совета Безопасности.

54. Несколько государств предложили инициировать специальный процесс, например, учредив для этого рабочую группу открытого состава. Было также высказано мнение о нецелесообразности инициирования нового процесса в рамках Организации Объединенных Наций в настоящее время. Было высказано мнение о том, что любой процесс по этому вопросу в рамках Организации Объединенных Наций должен основываться на консенсусе.

## **VII. Замечания и выводы Генерального секретаря**

55. ИИ способен повлиять на все сферы нашей жизни. Его применение в мирных целях может сыграть важную роль в содействии выполнению обязательств и достижению целей в области развития, включая цели в области устойчивого развития.

56. В военной области ИИ способен принести пользу как вооруженным силам, применяющим его, так и гражданскому населению, благодаря повышению точности в проведении операций и уменьшению вероятности человеческих ошибок. В то же время применение ИИ в военной области порождает серьезные проблемы, главная из которых заключается в сохранении ответственности и подотчетности человека.

<sup>1</sup> Более подробную информацию см. в документе [A/79/88](#).

57. Содержащееся в резолюции 79/239 Генеральной Ассамблеи заявление, подтверждающее, что нормы международного права, включая Устав Организации Объединенных Наций, международное гуманитарное право и международное право прав человека, применяется на всех этапах жизненного цикла ИИ, является важной отправной точкой. Вместе с тем важные вопросы о порядке их применения еще не решены.

58. Особого внимания требует использование военного ИИ в ситуациях, связанных с применением силы. Несмотря на потенциальные преимущества с точки зрения защиты гражданского населения и комбатантов, описываемые способы применения ИИ в современных конфликтах вызывают опасения в плане контроля со стороны человека и роли ИИ в поддержке боевых действий в густонаселенных районах. Машины, способные и уполномоченные по своему усмотрению лишать людей жизни, неприемлемы с политической точки зрения и морально отвратительны.

59. Риски, порождаемые ядерным оружием, сохраняются до тех пор, пока само это оружие не будет ликвидировано. В период до полной ликвидации ядерного оружия я настоятельно призываю все государства, обладающие таким оружием, согласиться с тем, что любое решение о применении ядерного оружия должно приниматься людьми, а не машинами.

60. ИИ может понизить порог для государственных и негосударственных субъектов на пути к разработке или приобретению химического и биологического оружия. Поэтому я настоятельно призываю государства в полной мере выполнять свои обязательства в рамках соответствующих рамочных документов по разоружению, нераспространению и контролю над вооружениями, а также систематически проводить оценку проблем и последствий, связанных с применением ИИ, для этих документов и хорошо подготовиться к реагированию на них.

61. Потенциальное внедрение гражданских приложений ИИ в военной области вызывает все большую озабоченность. В силу своего многоцелевого характера технологии ИИ создают проблемы для обеспечения надзора, транспарентности и подотчетности. Я настоятельно призываю государства внимательно изучить вопрос о размывании границ между новшествами в сфере гражданских приложений ИИ и их потенциальным применением в военной области.

62. Создание дополнительных механизмов сотрудничества в сфере ИИ, особенно на региональном и субрегиональном уровнях, представляется весьма целесообразным. Региональные и субрегиональные организации обладают уникальными возможностями для разработки и осуществления мер по обеспечению транспарентности и укреплению доверия в качестве средства уменьшения риска. Поэтому я призываю государства рассмотреть возможность разработки мер по обеспечению транспарентности и укреплению доверия на региональном и субрегиональном уровнях с учетом уникальных особенностей ИИ и связанных с ним проблем.

63. Под эгидой Организации Объединенных Наций ведутся всесторонние дискуссии по вопросу о мирном использовании и регулировании ИИ на благо человечества, в частности в контексте осуществления Глобального цифрового договора. Тем не менее, рассмотрение государствами-членами вопроса о применении ИИ в военной области в основном происходит за рамками форумов Организации Объединенных Наций. Резолюция 79/239 Генеральной Ассамблеи и настоящий доклад являются заметными первыми шагами, направленными на то, чтобы перенести эту важную дискуссию в Организацию Объединенных Наций. Я призываю государства проводить эти обсуждения инклюзивным и

конструктивным образом, с тем чтобы продвигать общее понимание и укреплять международное сотрудничество в целях уменьшения рисков.

64. Государствам рекомендуется изучить возможности принятия соответствующих мер, в том числе по наращиванию потенциала, для обеспечения конструктивного участия всех государств в процессах по этой теме, протекающих в Организации Объединенных Наций, что имеет существенно важное значение для укрепления общего понимания, разработки общих подходов и уменьшения потенциальных рисков.

65. Генеральная Ассамблея продемонстрировала свою способность утверждать мандаты на осуществление процессов, способствующих всестороннему обсуждению вопросов, касающихся новейших технологий и международной безопасности, а также задействовать заинтересованные стороны, включая международные и региональные организации, гражданское общество, научное сообщество и отраслевые круги. Такой подход с участием многих заинтересованных сторон особенно важен в сфере ИИ, в которой инновации исходят в основном от частного сектора, а значительная часть специалистов являются представителями академического и научного сообщества, а не правительственных кругов.

66. **Я рекомендую государствам изучить идеи, содержащиеся в настоящем докладе, и на восьмидесятой сессии Генеральной Ассамблеи предпринять конкретные шаги с целью организации целенаправленного и инклюзивного процесса для всестороннего рассмотрения вопроса о применении ИИ в военной области и его последствиях для международного мира и безопасности.**



## Приложение I

### Полученные ответы

#### А. Государства-члены

##### Аргентина

[Подлинный текст на испанском языке]  
[10 апреля 2025 года]

Настоящий доклад представляется в соответствии с резолюцией 79/239 Генеральной Ассамблеи от 24 декабря 2024 года, озаглавленной «Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности».

##### Общие замечания

Аргентинская Республика признает, что развитие искусственного интеллекта (ИИ) в военной области имеет значительные стратегические последствия. Его применение дает конкретные преимущества в ряде нелетальных функций, но в то же время порождает риски, требующие внимания в сфере международного права, этики и оперативной ответственности. В этой связи разработка и использование этих технологий должны осуществляться с соблюдением норм международного гуманитарного права и прав человека, а также при обеспечении постоянной ответственности и контроля за принятием важнейших решений со стороны человека.

##### Возможности

Военный ИИ, особенно применяемый для решения нелетальных задач, является законным и ценным инструментом наращивания национального потенциала. К первоочередным видам его применения относятся следующие:

- рационализация материально-технического снабжения и операционного обеспечения;
- поддержка обработки разведывательных данных;
- укрепление киберзащиты;
- составление моделей, обучение и стратегическое планирование.

Эти возможности способствуют повышению уровня эффективности и безопасности операций с учетом современных сценариев, повышая надежность обороны без ущерба для гуманитарных принципов и международных обязательств государства.

##### Проблемы

Ускоренное развитие военного ИИ порождает проблемы, которые необходимо решать совместными усилиями, в том числе следующие:

- понижение порога применения силы и сокращение времени для принятия решения человеком;
- вероятность необнаруженной алгоритмической предвзятости;
- распространение автономных систем среди негосударственных субъектов;
- риск расширения технологического разрыва между государствами.

Эти риски усиливают потребность в принятии общих принципов, поддающихся проверке гарантий и механизмов сотрудничества.

### **Регулирование, международное сотрудничество и инклюзивный подход в сфере технологий**

Мы считаем, что любой процесс регулирования в этой области должен быть построен на следующих принципах:

- недопущение принятия общих или преждевременных норм, ограничивающих автономное развитие законных военных технологий;
- четкое разграничение между летальными и нелетальными видами применения;
- обеспечение полноценного человеческого контроля в качестве важнейшего практического и нормативного условия;
- содействие всестороннему международному сотрудничеству в целях наращивания потенциала и сужения технологического разрыва между государствами.

Аргентина подтвердила эти принципы на недавних многосторонних форумах, подчеркнув важность продвижения общих норм ответственного применения ИИ в военной области, особенно в сфере киберзащиты и кибербезопасности.

В качестве примера региональной инициативы можно привести заседание рабочего комитета по вопросам ответственной разработки военного искусственного интеллекта, его применения и регулирования, которое прошло в 2024 году в Мендосе, Аргентина, в рамках XVI Конференции министров обороны стран Северной и Южной Америки и было посвящено совместной разработке международных стандартов.

### **Ссылка на Пакт во имя будущего**

В заключение следует отметить, что Аргентинская Республика официально отказалась признавать Пакт во имя будущего, упомянутый в преамбуле к резолюции [79/239](#) Генеральной Ассамблеи. Поэтому его упоминание не налагает на аргентинское государство никаких обязательств, не означает приверженности ему и не влечет за собой его поддержку.

### **Австрия**

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

В соответствии с просьбой, содержащейся в пункте 7 резолюции [79/239](#) Генеральной Ассамблеи, Австрия хотела бы поделиться нижеследующими соображениями и замечаниями, подготовленными в ее национальном качестве.

### **Использование искусственного интеллекта для обеспечения кибербезопасности и киберзащиты**

Программные средства обеспечения кибербезопасности, созданные на основе искусственного интеллекта (ИИ), уже широко используются для обнаружения вторжений и других вредоносных действий в компьютерных сетях. Такие средства ИИ, вероятно, обеспечат дальнейшую автоматизацию защиты информационно-вычислительных систем путем выявления факторов уязвимости и

подозрительных операций в целях повышения степени устойчивости аппаратно-программных комплексов.

В то же время средства ИИ все чаще используются для совершения все более изощренных кибератак и создания новых компьютерных вирусов в рамках соперничества между наступательными и оборонительными видами использования ИИ в сфере кибербезопасности. При этом программное обеспечение с поддержкой ИИ, включая большие языковые модели, упрощает злоумышленникам задачу совершения вторжений, поскольку они все чаще получают возможности создания вредоносных программ, не обладая обширными навыками программирования.

### **Использование искусственного интеллекта для проведения кампаний дезинформации в качестве элемента гибридных стратегий**

Программное обеспечение с поддержкой ИИ, способное создавать и распространять фальсифицированный контент, все чаще используется для повышения эффективности кампаний дезинформации. Для этого используются такие методы, как применение генеративного ИИ для создания персонализированного и локализованного контента в больших масштабах. Кроме того, быстро совершенствуется и уже широко используется программное обеспечение с поддержкой ИИ, предназначенное для синтеза глубоких фейков аудио и видео контента. Такой фальсифицированный контент может распространяться через массивные социальные бот-сети с поддержкой ИИ, чтобы создавать видимость изменения общественного мнения. Таким образом, ИИ упрощает задачу проведения масштабных кампаний дезинформации, поскольку количество и качество создаваемого фальшивого контента теперь не ограничивается количеством или квалификацией людей-операторов.

Вместе с тем алгоритмы ИИ могут использоваться и для выявления созданного ИИ контента и астротурфинговых кампаний, а псевдореальные аудио- и видео дипфейки проще всего выявлять с помощью специальных средств ИИ. Такие средства с поддержкой ИИ необходимо использовать для противодействия негативным последствиям использования ИИ в целях проведения кампаний дезинформации.

### **Использование искусственного интеллекта для распространения оружия**

ИИ может упростить приобретение оружия, в том числе оружия массового уничтожения. Большие языковые модели и созданные на их основе приложения, которые позволяют одним нажатием кнопки получать доступ к специальным знаниям, могут облегчить злоумышленникам задачу изготовления оружия. ИИ может использоваться в самых различных целях: от получения доступа к чертежам или печатания компонентов стрелкового оружия и легких вооружений до модификации патогенов, предназначенных для использования в качестве биологических средств поражения. Если наличие легко доступных знаний приводит к сокращению сферы охвата и уменьшению масштабов оружейных программ, то это усложнит задачу обнаружения таких угроз, их предотвращения и подготовки к ним.

В то же время алгоритмы машинного обучения можно использовать и для борьбы с распространением оружия. Они позволяют обнаруживать аномалии и распознавать образы, что может способствовать выявлению вредоносной деятельности, в том числе путем обнаружения незаконных денежных потоков для оружейных программ или путем анализа закономерностей в спутниковых данных.

### **Использование искусственного интеллекта для осуществления проверки в сфере контроля над вооружениями и принятия решений в кризисных ситуациях**

ИИ может способствовать осуществлению проверки соглашений о контроле над вооружениями. Это объясняется тем, что он способен анализировать большие объемы данных — например, из таких источников, как спутниковые изображения — и классифицировать различные объекты. Благодаря этому он может идентифицировать военную технику, например танки и ракеты, обнаруживать казармы или выявлять такую военную деятельность, как передвижение войск и проведение учений. Кроме того, как уже говорилось, ИИ может способствовать выявлению незаконных оружейных программ. Это весьма осложнит задачу тех, кто пытается нарушать соглашения о контроле над вооружениями, а государства-участники могли бы быть уверены в том, что все соблюдают положения таких соглашений.

Увеличение объема и повышение качества информации благодаря способности ИИ анализировать и классифицировать данные, поступающие с датчиков, могут не только облегчить выполнение соглашений о контроле над вооружениями, но и способствовать принятию более рациональных решений в ситуациях, когда в отношениях между государствами особенно обостряется военная напряженность. Политические и военные лидеры могли бы воспользоваться возможностями ИИ для улучшения ситуационной осведомленности в целях деэскалации кризисов.

### **Мир и безопасность и Устав Организации Объединенных Наций**

Одна из специфических проблем, связанных с применением ИИ в военной области, заключается в потенциальном риске для мира и безопасности вследствие непреднамеренной эскалации и недоразумений, которые могут возникнуть при использовании ИИ. Использование средств машинного обучения создает дополнительный уровень сложности, поскольку не все субъекты способны до конца понимать, как работает та или иная система.

Кроме того, для использования ИИ в системах поддержки принятия решений в отношении взаимодействия человека и машины и для определения необходимого участия человека необходимо принять меры и внедрить защитные механизмы в целях обеспечения подотчетности и ответственности, а также для смягчения алгоритмической предвзятости.

Все эти риски необходимо уменьшать путем обеспечения надзора и принятия мер, учитывающих специфические проблемы, возникающие при использовании таких технологий.

Следует отметить, что статьей 36 Дополнительного протокола к Женевским конвенциям от 12 августа 1949 года, касающегося защиты жертв международных вооруженных конфликтов (Протокол I), предусматривается обязательство изучать законность всех новых видов оружия, средств или методов ведения войны до их применения в вооруженном конфликте.

Кроме того, ИИ может применяться для поддержки эффективного выполнения обязательств по международному гуманитарному праву, в частности, когда речь идет о защите гражданских лиц, в качестве позитивного обязательства и позитивных действий, в том числе по линии проектов, исследовательской работы и создания приложений, специально предназначенных для решения этой задачи.

## **Рамочные основы международного сотрудничества и обмена информацией**

Поскольку проблема, связанная с применением ИИ в военной области, все более усложняется и ставит сложные задачи перед всеми государствами, весьма актуальное значение приобретают многосторонние обсуждения и форматы для обмена опытом и примерами передовой практики. В этой связи Австрия поддержала Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях. В качестве сопредседателя рабочей группы по надзору за осуществлением указанной декларации Австрия вместе с Германией способствует обмену передовым опытом в решении проблем и разработке политики в этой области. Кроме того, Австрия поддержала План действий Саммита по ответственному использованию искусственного интеллекта в военной сфере, а также Парижскую декларацию о сохранении человеческого контроля в оружейных системах с поддержкой ИИ.

## **Связь между работой международного сообщества над вопросом о применении искусственного интеллекта в военной области и его работой над вопросом о смертоносных автономных системах вооружений**

Что касается сферы применения ИИ и автономных систем в военной области в целом, то здесь следует выделить автономные системы вооружений. Автономные системы вооружений вызывают особую озабоченность с точки зрения права, этики и безопасности. Этот вопрос не рассматривается непосредственно в резолюции [79/239](#) Генеральной Ассамблеи, поскольку соответствующие дискуссии в рамках Организации Объединенных Наций ведутся уже с 2013 года, и все большее число государств выражают желание установить правила и ограничения в отношении автономных систем вооружений на международном уровне. Поэтому в своих комментариях для настоящего доклада Австрия хотела бы лишь особо отметить, что она выступает за разработку юридически обязывающего документа относительно автономных систем вооружений, и отметить важную работу, проводимую в настоящее время Группой правительственных экспертов в рамках Конвенции о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие, а также отметить дополнительные усилия, предпринятые в рамках самой первой резолюции Генеральной Ассамблеи по смертоносным автономным системам вооружений (резолюция [78/241](#)), в соответствии с которой был опубликован доклад Генерального секретаря ([A/79/88](#)), и последующей резолюции (резолюция [79/62](#)), в соответствии с которой было принято решение созвать 12 и 13 мая 2025 года в Нью-Йорке неофициальные консультации по вопросу о смертоносных автономных системах вооружений.

## **Соображения, касающиеся рамок основ правового регулирования искусственного интеллекта**

Закон Европейского союза об искусственном интеллекте определяет законодательную базу Европейского союза для систем ИИ в различных секторах и имеет целью укрепление доверия к приложениям с использованием ИИ и содействие использованию преимуществ ИИ с соблюдением прав человека, основных свобод и демократических ценностей. Этот закон подчеркивает важность транспарентности, подотчетности и контроля со стороны человека при разработке и внедрении систем ИИ, способствуя при этом обеспечению правовой определенности, новаторства и конкурентоспособности. Этот закон не распространяется на системы ИИ, разработанные для военных и оборонных целей или для обеспечения национальной безопасности. Вместе с тем в Законе об ИИ применяется подход, основанный на оценке рисков, который может принести пользу при

рассмотрении широкого спектра потенциальных применений ИИ в военной области.

### **Перспективы**

Австрия высоко оценивает работу, проводимую в различных упомянутых выше форматах и форумах по вопросам применения ИИ в военной сфере, и убеждена в том, что они будут способствовать формированию свода согласованных на международном уровне норм и стандартов, обеспечивающих ответственное применение ИИ в военной области в соответствии с международно-правовыми обязательствами и этическими принципами.

### **Чили**

[Подлинный текст на испанском языке]  
[11 апреля 2025 года]

Ранее Чили уже отмечала, что стремительное развитие новых и новейших технологий является актуальным вопросом международной безопасности и является вызовом для всех стран. Эти новые технологии, в частности искусственный интеллект (ИИ), могут принести огромную пользу для развития и благосостояния общества, но в то же время поднимают важные вопросы, касающиеся последствий их применения для обеспечения безопасности и обороны. Применение новых технологий может принести значительные преимущества, но при этом порождает риски и трудности.

В этой связи Чили считает желательным выработать общее понимание концепции ответственного применения ИИ в военной области и в сфере безопасности, а также разработки и применения так называемых смертоносных автономных систем вооружений. Чили поддерживает многосторонние усилия по налаживанию и укреплению диалога и дискуссий между странами в целях достижения консенсуса и взаимопонимания по вопросам применения этих новых технологий.

Чили занимает лидирующие позиции в сфере искусственного интеллекта благодаря значительным достижениям в создании благоприятных условий для внедрения этой технологии, а также новаторству в сфере политики и регулирования ИИ. В октябре 2021 года в Чили была принята первая Национальная стратегия в сфере ИИ, разработанная в сотрудничестве с различными государственными и частными структурами. В этой стратегии основной упор делается на трех ключевых областях: создание благоприятных условий, применение и развитие технологии, а также создание нормативно-правовой и этической базы для обеспечения ответственного и безопасного применения ИИ.

В 2024 году Чили приняла обновленную версию Национальной стратегии в сфере ИИ, включающую такие новые дополнительные направления, как международное взаимодействие, окружающая среда и климатический кризис, инклюзивный подход и запрещение дискриминации, дети и подростки, а также культура и сохранение культурного наследия. Эта стратегия дополняется соответствующим планом действий, который предусматривает более 100 мероприятий на 2026 год и охватывает такие сферы, как образование, здравоохранение, окружающая среда и культура. Кроме того, новая Национальная стратегия в сфере ИИ включает принципы, сформулированные в Рекомендации Организации Объединенных Наций по вопросам образования, науки и культуры (ЮНЕСКО) об этических аспектах искусственного интеллекта, что обеспечивает ее соответствие самым современным международным рамкам.

Следует отметить, что Чили стала первой страной в мире, заполнившей формуляр оценки готовности, разработанный ЮНЕСКО для определения степени готовности страны к внедрению ИИ на этической и ответственной основе, подтвердив свою приверженность выполнению Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта на национальном политическом уровне. Чили выступает за этическое и ответственное развитие технологий, о чем свидетельствует участие страны в саммитах по ИИ, организованных Соединенным Королевством (2023 год), Республикой Корея (2024 год) и Францией (2025 год).

На законодательном уровне в Чили в настоящее время обсуждается законопроект о регулировании систем ИИ на основе оценки рисков с целью содействия разработке и внедрению ИИ в условиях соблюдения демократических принципов и основных прав человека.

Что касается обороны и безопасности, то Чили поддержала проведение и приняла активное участие в работе саммитов по ответственному применению искусственного интеллекта в военной сфере, которые прошли в Гааге (2023 год) и Сеуле (2024 год). Наша страна одобрила итоговые документы обоих саммитов («Призыв к действиям» (2023 год) и «План действий» (2024 год)). Кроме того, Чили поддерживает работу Глобальной комиссии по ответственному применению искусственного интеллекта в военной области.

Следует также отметить, что 13–14 июня 2024 года в Чили состоялся региональный практикум по ответственному применению искусственного интеллекта в военной области и сфере безопасности в целом, организованный Министерством иностранных дел Чили и Министерством иностранных дел Коста-Рики под эгидой Королевства Нидерландов и Республики Корея. Кроме того, это мероприятие получило поддержку Центра правовых, технологических и общественных исследований при Чилийском университете и Женевского центра за гуманитарный диалог, которые также приняли участие в его организации. В практикуме приняли участие представители Аргентины, Бразилии, Доминиканской Республики, Колумбии, Коста-Рики, Мексики, Парагвая, Сальвадора, Тринидада и Тобаго, Уругвая, Чили, Эквадора, Ямайки, а также Королевства Нидерландов и Кореи. Со стороны Чили в нем также приняли участие представители Министерства национальной обороны и вооруженных сил.

Чили считает, что применение ИИ в военной области и в сфере обеспечения безопасности может открывать такие возможности и преимущества, как совершенствование процесса принятия решений и стратегического анализа, повышение эффективности материально-технического снабжения, наращивание потенциала киберзащиты и кибербезопасности путем усиления безопасности критической инфраструктуры, а также способствовать планированию сложных миссий по поддержанию мира и оказанию гуманитарной помощи. Кроме того, оно может способствовать наращиванию потенциала проверки и мониторинга в области контроля над вооружениями и соблюдения режимов контроля над вооружениями.

Чили считает, что технологии ИИ необходимо разрабатывать, внедрять и применять в соответствии с нормами международного права, включая, в соответствующих случаях, Устав Организации Объединенных Наций, нормы международного гуманитарного права, международного права прав человека и другие соответствующие правовые основы.

Чили считает важным принимать такие меры обеспечения контроля и безопасности, которые не позволяли бы безответственным субъектам приобретать и неправомерно использовать потенциально опасные средства ИИ в военной

области, включая системы с поддержкой ИИ, обеспечивая при этом, чтобы такие меры не ограничивали равный доступ к преимуществам средств ИИ в других невоенных областях.

Кроме того, Чили считает важным объединить усилия по предотвращению использования технологий ИИ для содействия распространению оружия массового уничтожения государственными и негосударственными субъектами, включая террористические группы, и подчеркнуть, что технологии ИИ должны поддерживать усилия в области разоружения, контроля над вооружениями и нераспространения, а не препятствовать им. Особенно важно сохранить контроль и участие человека во всех действиях, имеющих критически важное значение для информационного обеспечения и осуществления суверенных решений, касающихся применения ядерного оружия, без ущерба для конечной цели, которая заключается в построении мира, свободного от ядерного оружия.

Чили выступает за разработку таких мер укрепления доверия, как обмен информацией и проведение консультаций по вопросам передового опыта и извлеченных уроков между государствами. В этой связи Чили считает важным, чтобы страны могли разработать и установить национальные стратегии, принципы, нормы, политику, рамки и законодательство для обеспечения ответственного применения ИИ в военной области. Меры по укреплению доверия, способствующие повышению степени транспарентности, могут стать эффективным инструментом для разработки механизмов сдерживания, контроля и обеспечения доверия как на национальном, так и на международном уровне.

Кроме того, Чили считает крайне важным сократить отставание в сфере цифровых технологий и ИИ между развитыми и развивающимися странами и необходимым повысить уровень понимания и осведомленности о последствиях применения искусственного интеллекта в военной области, в том числе путем обмена знаниями и передовым опытом между всеми государствами.

В этой связи Чили считает необходимым разрабатывать инициативы и программы, способствующие наращиванию потенциала, особенно в развивающихся странах, чтобы содействовать полноценному участию этих стран в дискуссиях по вопросам управления ИИ в военной области, при том понимании что такое наращивание потенциала также может помочь странам более глубоко вникнуть в проблему применения ИИ в военной области и способствовать развитию, внедрению и применению военных средств ИИ ответственным и законным образом. Наращивание потенциала также позволит странам лучше подготовиться к международным дискуссиям и диалогам.

Чили придает большое значение укреплению международного сотрудничества в сфере наращивания потенциала путем поощрения диалога и дискуссий на национальном, региональном, субрегиональном и межрегиональном уровнях, включая учебные программы, конференции, практикумы и семинары, в том числе для должностных лиц, занимающихся дипломатическими, политическими и техническими вопросами, с тем чтобы восполнить пробел в знаниях об ответственной разработке, внедрении и применении ИИ в военной области.

Чили ценит региональные и субрегиональные дискуссии и диалог по вопросам применения ИИ в военной области и считает важным поощрять, особо отмечая в этой связи XVI Конференцию министров обороны стран Северной и Южной Америки, которая прошла в Аргентине 13–16 октября 2024 года, и, в частности, принятую на этой конференции Мендосскую декларацию, в которой, среди прочего, рекомендуется поощрять этическое применение ИИ в военной области, учитывать экономическое и технологическое многообразие стран — участниц Конференции, а также содействовать развитию механизмов



укрепления взаимного доверия и сотрудничества на полушарии и в регионе, которые позволят странам — участницам Конференции обмениваться знаниями и передовым опытом, разрабатывать согласованные стандарты и наращивать технологический потенциал в целях применения ИИ в военной области.

Наконец, Чили считает, что к обсуждению и диалогу по вопросам применения ИИ в военной области необходимо привлекать все заинтересованные стороны, в частности гражданское общество, научные круги, отраслевые структуры, частный сектор, техническое сообщество, региональные и международные организации.

## Китай

[Подлинный текст на китайском языке]  
[11 апреля 2025 года]

Быстрое развитие и широкое применение искусственного интеллекта (ИИ) в военной области изменяет парадигмы будущих войн и создает потенциальные проблемы для международного мира и безопасности. Поскольку мир сталкивается с многочисленными вызовами миру и безопасности, все стороны должны стремиться к поиску консенсуса на основе диалога и сотрудничества по вопросам регулирования военного применения ИИ, содействовать разработке открытой, справедливой и эффективной системы управления безопасностью ИИ и сводить к минимуму риски, чтобы сохранять безопасность, надежность и контролируемость технологий ИИ и обеспечивать их развитие исключительно в интересах прогресса человеческой цивилизации.

Китай всегда ответственно и конструктивно участвовал в решении вопросов глобального управления военными применениями ИИ. Мы выступаем за то, чтобы придерживаться концепции «человекоцентричного подхода к военным применениям ИИ» и поддерживать принцип обеспечения общей, всеобъемлющей, совместной и устойчивой безопасности в стремлении к построению сообщества единой судьбы человечества. В 2021 году в рамках Конвенции по конкретным видам обычного оружия Китай представил документ с изложением позиции по вопросам регулирования военного применения ИИ, в котором были предложены систематизированные мнения и рекомендации касательно ответственного развития и применения ИИ в военной области с точки зрения стратегической безопасности, военной политики, права и этики, технологической безопасности, проведения научно-исследовательских и опытно-конструкторских работ, управления рисками и осуществления контроля, разработки правил и обеспечения международного сотрудничества. В 2023 году Китай выступил с Инициативой глобального управления ИИ, призвав все страны, особенно крупные державы, принять взвешенный и ответственный подход к проведению научно-исследовательских и опытно-конструкторских работ и применению технологий ИИ в военной области. Наши конкретные предложения включают следующее:

Во-первых, необходимо принять взвешенный и ответственный подход. В процессе наращивания своего законного национального оборонного потенциала все страны, особенно крупные державы, должны воздерживаться от стремления к получению абсолютного военного превосходства с помощью ИИ и подрыва законных интересов безопасности других стран. Необходимо прилагать усилия к тому, чтобы избегать недоразумений и просчетов и не допускать гонки вооружений в этой области.

Во-вторых, необходимо придерживаться человекоцентричной концепции. Существенно важно всегда рассматривать человека как конечного субъекта ответственности и обеспечивать, чтобы соответствующие системы вооружений были подконтрольны человеку. Военное применение ИИ должно основываться на уважении и защите человеческого достоинства и прав человека, а также на уважении общих ценностей человечества.

В-третьих, необходимо соблюдать основной принцип «ИИ во благо». Применение ИИ в военной области должно способствовать поддержанию мира, соответствовать нормам международного гуманитарного права и другим применимым нормам международного права, а также поддерживать цель уменьшения количества сопутствующих жертв.

В-четвертых, необходимо осуществлять гибкое управление. Мы должны усилить перспективную оценку рисков и подготовку специалистов по ИИ, принимать необходимые меры по смягчению рисков и уменьшать риски распространения, не препятствуя при этом новаторству и мирному использованию технологий.

В-пятых, необходимо сохранять многосторонний подход. Мы должны поддерживать Организацию Объединенных Наций в ее надлежащей роли, приветствовать создание инклюзивных платформ для обсуждения с участием всех сторон и прилагать усилия для создания механизмов регулирования, основанных на всеобщем участии и широком консенсусе.

Китай считает, что значению ИИ в военной области следует дать объективную оценку. Важно направлять развитие военного ИИ в нужное русло, не допуская его выхода из-под регулирующего контроля. На следующем этапе международное сообщество должно сотрудничать, чтобы извлекать максимальную пользу и сводить вред к минимуму. Китай предлагает следующие идеи и предложения:

Во-первых, нужно установить четкие правила. Вопросам безопасности и развития необходимо уделять одинаковое внимание. Настоятельно необходимо придерживаться целей и принципов Устава Организации Объединенных Наций, соблюдать основные нормы, регулирующие международные отношения, и следить за тем, чтобы технология ИИ не стала инструментом для осуществления вторжений в другие страны и достижения гегемонии. Китай готов к дальнейшему обмену мнениями со всеми сторонами относительно человекоцентричной концепции в сфере военного применения ИИ и последовательного стремления к достижению консенсуса.

Во-вторых, нужно усовершенствовать меры регулирования. С учетом текущего положения дел в сфере разработки и применения ИИ мы должны содействовать созданию системы тестирования и оценки, осуществлять гибкое регулирование и внедрять ступенчатую классификацию управления в целях обеспечения оперативного и эффективного реагирования. Все страны должны, исходя из своих национальных условий, принимать и совершенствовать внутренние законы и нормативные акты, совершенствовать соответствующие этические принципы, а также укреплять систему образования и подготовки кадров в целях повышения уровня безопасности, надежности и управляемости технологий ИИ.

В-третьих, нужно укреплять международное сотрудничество. Все страны должны придерживаться принципов открытости и инклюзивности, вести диалог и обмениваться мнениями для углубления взаимопонимания, а также укреплять координацию политики и сотрудничество по вопросам наращивания потенциала в сфере управления ИИ, чтобы постоянно совершенствовать возможности управления.

## Египет

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

В соответствии с резолюцией 79/239 Генеральной Ассамблеи правительство Арабской Республики Египет хотело бы поделиться своим мнением о возможностях и проблемах, которые создает применение искусственного интеллекта (ИИ) в военной области для международного мира и безопасности.

Резолюция 79/239 Генеральной Ассамблеи представляет собой важный шаг на пути укрепления многостороннего подхода к теме применения ИИ в военной области и повышения ее значимости в политической повестке дня. Это произошло после того, как Генеральный секретарь выступил с призывом разработать нормы, правила и принципы проектирования, разработки и применения средств ИИ в военной области с привлечением всех заинтересованных сторон.

Понимая, что в указанной резолюции, в соответствии с которой представляются настоящие соображения, особый упор делается не на смертоносных автономных системах вооружений, мы считаем необходимым повторить неизменную позицию Египта, согласно которой при любом содержательном обсуждении этого вопроса нужно в первую очередь рассматривать все соображения этики, права и обеспечения безопасности, связанные со смертоносными автономными системами вооружений, которые представляют собой наиболее серьезную угрозу для поддержания международного мира и безопасности в том, что касается военного применения ИИ.

Как уже предлагал Генеральный секретарь, наиболее эффективным и реалистичным подходом является заключение соглашения относительно юридически обязывающего запрета на смертоносные автономные системы вооружений, которые функционируют без человеческого контроля или надзора и не могут применяться в соответствии с нормами международного гуманитарного права. Применение двухуровневого подхода, включающего запрет и ограничение и/или регулирование, в том числе запрет систем вооружений, функционирующих без человеческого контроля, и регулирование других систем, имеет существенно важное значение для создания необходимой универсальной правовой архитектуры, которая обеспечила бы благоприятные условия для получения максимальной пользы от использования новых возможностей, открывающихся благодаря применению ИИ в военной области, при одновременном решении соответствующих проблем реалистичным, эффективным и своевременным образом.

Международная политика в вопросе применения ИИ в военной области далеко не единообразна. Египет внимательно следит за многочисленными международными инициативами в этой области, которые свидетельствуют о все более широком осознании сопутствующих рисков. Тем не менее, обсуждение этих инициатив выявило расхождения во взглядах, восприятии угроз и представлении об очередности задач, поэтому мы должны предостеречь от опасности создания фрагментированной политической структуры или инициирования конкурирующих процессов, как это было в связи с другими сферами применения новых и новейших технологий.

Эти инициативы, безусловно, необходимо упорядочить и объединить под эгидой Организации Объединенных Наций, с тем чтобы обеспечить их инклюзивность и эффективность. Организация Объединенных Наций и ее разоруженческий механизм представляют собой единственную эффективную и всеобъемлющую платформу для разработки необходимых международных правил и нормативной базы, особенно с учетом того, что развитие технологий по-прежнему

значительно опережает усилия по обеспечению необходимого регулирования на международном уровне.

Поэтому крайне важно создать под эгидой Организации Объединенных Наций универсальную, независимую, единую и надежную платформу для обсуждения вопросов будущего управления ИИ в военной области. Предлагаемый процесс под руководством Организации Объединенных Наций следует организовать таким образом, чтобы избежать некоторых уже наметившихся контрпродуктивных расхождений в подходах. Одно из таких расхождений заключается в несогласованности законных усилий по обеспечению соблюдения правовых и этических норм и преследованием военных интересов без должного учета гуманитарных последствий.

Следует также подчеркнуть, что мы высоко оцениваем обсуждения, проведенные в рамках Группы правительственных экспертов по смертоносным автономным системам вооружений под эгидой Конвенции о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие, однако считаем, что эта платформа не может заменить собой намеченный в рамках Организации Объединенных Наций процесс по вопросам применения ИИ в военной области, поскольку эта группа правительственных экспертов не является универсальной по своему характеру и не имеет мандата для рассмотрения столь многоплановой и своеобразной темы. Кроме того, вызывает сожаление тот факт, что Группа правительственных экспертов достигла лишь минимального прогресса и пока не добилась каких-либо конкретных результатов.

Технологии ИИ не только открывают новые возможности, но и порождают множество рисков, связанных с особенностями таких технологий, которые могут функционировать непредсказуемым и необъяснимым образом. Эти риски включают в частности распространение дезинформации, угрозу непреднамеренной эскалации и создание киберрисков, а также нецелевое использование и пространственную деятельность со стороны негосударственных субъектов. Такие риски могут быть новыми или могут усложнять существующие риски.

Общепризнанно, что существует широкий спектр возможных военных применений ИИ. Вместе с тем серьезные усилия по разработке будущего механизма управления ими должны определить правильный порядок очередности задач с учетом создаваемого риска и последствий для мира и безопасности. Это позволит обеспечить проведение целенаправленного и упорядоченного обсуждения и не допускать неоправданного рассеивания внимания. При этом Египет твердо придерживается мнения о том, что помимо смертоносных автономных систем вооружений необходимо сделать упор и на других возможностях автономных или полуавтономных систем, которые позволяют применять силу и/или понижают порог ее применения и, соответственно, могут активизировать гонку вооружений, охватывающую как обычные вооружения, так и оружие массового уничтожения. Потенциал повышения степени автономности ядерного оружия и таких современных обычных вооружений, как гиперзвуковые ракеты, чреват неизвестными рисками и может непредсказуемым образом изменить картину будущего конфликта.

Кроме того, нужно уделять особое внимание не столько планированию работы тыла, разведке, наблюдению и рекогносцировке, поскольку они не оказывают столь разрушительного воздействия, сколько командованию и управлению и выбору целей. Помимо этого, больше внимания следует уделять наступательным, а не оборонительным средствам.

Планируемые обсуждения в рамках намеченного процесса под руководством Организации Объединенных Наций должны быть нацелены прежде всего на достижение общего понимания основных элементов, лежащих в основе разработки, внедрения и применения ИИ в военной области. К числу таких элементов относятся:

- Всестороннее соблюдение норм применимого международного права, включая такие основополагающие принципы международного гуманитарного права, как необходимость, соразмерность и избирательность, а также учет других этических соображений на протяжении всего жизненного цикла и на всех этапах применения средств ИИ в военной области.
- Первостепенная важность сохранения участия человеческого элемента на протяжении всего жизненного цикла военных средств ИИ, включая участие в принятии решений, вмешательство, надзор и контроль со стороны человека в качестве ключевых факторов обеспечения ответственности. Необходимо обеспечить, чтобы все программное обеспечение, алгоритмы и проектные разработки, связанные с применением средств ИИ в военной области, проходили критический анализ с участием человека и соответствовали принципу объяснимости. Хотя правительства утверждают, что доктриной предусматривается сохранение контроля со стороны человека над системами, использующими ИИ, у некоторых из них может возникнуть соблазн в угоду военным интересам сделать свои системы вооружений более автономными.
- Баланс между уменьшением рисков распространения среди негосударственных субъектов и сдерживанием злонамеренного использования и сохранением прав государств на приобретение ИИ и технологий двойного назначения. Крайне важно избегать создания любых механизмов произвольного международного надзора или введения дискриминационного экспортного контроля в любой форме.
- Компонент, связанный с наращиванием потенциала с целью обеспечения надлежащих инвестиций в человеческий капитал, передачи технологий и обмена знаниями и передовым опытом таким образом, чтобы сохранить право развивающихся стран на получение потенциальных выгод от различных видов военного применения ИИ, а также с целью преодоления отставания в сфере цифровых технологий.
- Границы применения ИИ в военной области и его взаимодействие с другими новыми и новейшими технологиями. Целесообразно обсудить пути обеспечения взаимодополняемости с другими процессами, протекающими под руководством Организации Объединенных Наций, включая Рабочую группу открытого состава по вопросам безопасности в сфере использования информационно-коммуникационных технологий и самих информационно-коммуникационных технологий, учитывая, например, взаимосвязь между ИИ и кибероперациями. Кроме того, в ходе обсуждений следует уделять внимание главным образом военной области, а не на более широком аспектам безопасности.

Наконец, важно обеспечить инклюзивность и сбалансированность при разработке путей управления в целях обеспечения ответственного, подотчетного и человекоцентричного подхода к ИИ в рамках многосторонних перспектив в Организации Объединенных Наций, чтобы внести ключевой вклад в обсуждение вопросов политики. Однако участие сторон не должно наносить ущерб суверенным прерогативам государств в вопросах выработки политики.

## Сальвадор

[Подлинный текст на испанском языке]  
[10 апреля 2025 года]

### Справочная информация

В последние годы применение искусственного интеллекта (ИИ) в военной области стало играть важную роль. В многочисленных докладах говорится, что эти новые технологии получают все большее развитие и распространение, что позволяет использовать такие вычислительные средства в процессах военного планирования и принятия решений, в том числе в вопросах определения целей. В связи с этим возникает множество вопросов относительно общих последствий, правовых последствий и рисков для гражданского населения, порождаемых применением этих технологий. Примером служат дебаты о последствиях применения этих технологий, проходящие в ходе многосторонних переговоров, на которых автономные системы вооружений рассматриваются с политической, правовой и гуманитарной точек зрения. Вместе с тем признается, что ИИ имеет гораздо более широкое военное применение.

Поэтому необходимо расширить наше понимание использования и применения ИИ в военной области, особенно для решения конкретных задач выбора военных целей и применения силы.

Вопрос ответственного применения ИИ в военной области приобрел особую актуальность благодаря дискуссиям на первом саммите по ответственному использованию искусственного интеллекта в военной сфере, который состоялся в Королевстве Нидерландов в феврале 2023 года. Кроме того, этому вопросу уже уделяется особое внимание в ходе заседаний Группы правительственных экспертов по смертоносным автономным системам вооружений, базирующейся в Женеве.

Следует отметить, что до сих пор вопросы применения и использования ИИ обсуждались в основном в рамках дискуссий об автономных системах вооружений, однако вопрос о применении ИИ в военной области стоит гораздо шире и приобретает новое измерение, особенно в отношении приложений, решающих задачу автоматизации некоторых военных функций, помимо тех, которые имеют целью обеспечение автономности систем вооружений.

В целом, ИИ — это новая тема, обсуждение которой все еще продолжается и развивается весьма быстрыми темпами. В этой связи на национальном, региональном и многостороннем уровнях разрабатываются инициативы, которые позволят учитывать его последствия. Очевидно, что страны Латинской Америки и Карибского бассейна не находятся на одном с развитыми странами уровне развития технологий и потенциала, которые позволят выявлять и понимать возможности и риски, возникающие в процессе применения ИИ. Поэтому важно определить национальную позицию, которая позволит им активно участвовать в дискуссиях, проходящих на международных форумах, и таким образом наладить сотрудничество в сфере наращивания потенциала и в конкретных аспектах, которое позволит им принимать активное участие в рассмотрении этой проблемы и понимать возможности и потенциальные риски с точки зрения безопасности на национальном, региональном и глобальном уровнях.

### Инициативы, в которых участвовал Сальвадор

- Сальвадор принял участие в саммите по ответственному использованию искусственного интеллекта в военной сфере, который состоялся в

2023 году в Королевстве Нидерландов, и присоединился к принятой по его итогам декларации (февраль 2023 года).

- Он принял участие в Латиноамериканской и Карибской конференции по социальным и гуманитарным последствиям, связанным с автономными вооружениями, на которой было принято Беленское коммюнике (февраль 2023 года).
- Сальвадор входит в «Группу шестнадцати» в рамках дискуссий Группы экспертов по смертоносным автономным системам вооружений. Хотя этот вопрос стоит обособленно, он имеет отношение к применению ИИ в военной области.

### **Национальная позиция**

- Некоторые приложения ИИ могут принести определенную пользу в военной области, особенно если они не связаны с функциями идентификации и распознавания военных целей и не имеют непосредственного отношения к применению силы, которая создает угрозу для гражданского населения, а связаны с решением административных задач, например анализом данных и автоматизированным обучением без участия человека в рамках военных операций.
- Вместе с тем неправомерное использование этих приложений может вызывать негативные последствия, особенно в плане защиты гражданского населения и объектов гражданской инфраструктуры, которые пользуются особыми режимами защиты в соответствии с нормами международного права, включая международное гуманитарное право и международное право прав человека.
- Важно применять подход, основанный на оценке рисков, который позволяет регулировать и запрещать некоторые функции ИИ, особенно те, которые ограничивают значимый контроль человека над применением силы, те, которые закрепляют алгоритмическую предвзятость вследствие использования нерепрезентативных баз данных или исторических данных и создают риски для прав человека, и те, которые в долгосрочной перспективе создают риски для международной безопасности, особенно когда машине дается право принимать решения, касающиеся жизни и смерти человека, или когда эти инструменты включают очень сложные технологические функции, такие как самообучение, что может иметь серьезные гуманитарные, социальные, экономические, политические и даже экологические последствия.
- В настоящее время настоятельно необходимо обеспечить надлежащее регулирование в сфере ИИ, поскольку оно имеет существенно важное значение для его безопасного и этического развития, что позволит защитить пользователей и общество от потенциальных злоупотреблений и рисков, а также будет способствовать поиску новаторских решений, обеспечивая четкую определенную и безопасную среду для разработчиков и исследователей.
- Хотя конечная цель заключается в разработке юридически обязывающих документов, считается, что развитие этих технологий происходит быстрее, чем развитие или разработка международного права в этой области, поэтому мы считаем целесообразным сохранить подход, в котором упор делается на ответственном поведении в качестве потенциальной основы для принятия всеобъемлющих правовых обязательств в целях более эффективного решения этой проблемы.

- Важно учитывать проблемы, которые возникают в связи с появлением новых технологий в сфере безопасности, например технологий материалов, таких как трехмерная печать, используемых в изготовлении стрелкового оружия и легких вооружений, применениями робототехники в создании автономных систем вооружений, а также некоторых видов использования и применения ИИ, который в силу своего двойственного характера может усиливать предвзятость в осуществлении функций командования и управления в вооруженных конфликтах, особенно создавая повышенный риск для гражданского населения.
- Утрата или подмена контроля в военной области может привести к непредвиденным рискам. ИИ может дополнять возможности человека, однако отсутствие контроля в военном контексте может порождать и другие риски, которые необходимо тщательно изучить. Поддержка со стороны ИИ в военной области должна подкреплять процесс принятия решений в конкретных ситуациях или обеспечивать его информационную поддержку, но ни в коем случае не подменять участие человека или человеческое суждение.
- Применение ИИ в военной области должно соответствовать нормам международного права, международного права прав человека и международного гуманитарного права и служить общественному благу.
- Необходимо наращивать потенциал стран, чтобы они могли выявлять риски, возникающие в результате неправомерного применения ИИ, и определять соответствие такого применения нормам международного права.
- Другие участники процесса создания и развития технологий такого рода, такие как предпринимательские и научные круги, должны быть вовлечены в многосторонние дискуссии, а международное сотрудничество между заинтересованными сторонами должно поощряться, чтобы использовать преимущества, которые дает использование ИИ в мирных целях, в интересах развития стран.

## Финляндия

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Финляндия с удовлетворением представляет свои мнения в соответствии с резолюцией 79/239 Генеральной Ассамблеи от 24 декабря 2024 года о применении искусственного интеллекта (ИИ) в военной области и его последствиях для международного мира и безопасности, в которой Ассамблея просила Генерального секретаря запросить мнения государств-членов о «возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений».

Принятие международных принципов или правил, регулирующих вопросы применения ИИ в военной области, имеет фундаментальное значение для обеспечения соблюдения норм международного права, укрепления безопасности и снижения потенциальных рисков возникновения конфликтов. В то же время необходимо создать условия для наращивания национального оборонного потенциала, который соответствовал бы нормам международного права. Финляндия взяла на себя обязательство ответственным образом подходить к разработке, внедрению и применению средств ИИ в военной области, руководствуясь нормами международного права, в частности международного гуманитарного



права, и действуя таким образом, чтобы не подрывать международный мир, безопасность и стабильность, одновременно продолжая научно-исследовательские и опытно-конструкторские работы и эксперименты, а также инновационные усилия в сфере технологий ИИ.

Все более важное значение приобретает определение последствий применения революционных технологий для политики в сфере внешних сношений, безопасности и обороны и разработка способов устранения таких последствий. Финляндия активно участвует в глобальных дискуссиях по вопросам регулирования технологий, отстаивая основные права и права человека и устраняя связанные с этим риски при разработке и применении ИИ и соответствующей политики.

Помимо выявления рисков, связанных с революционными технологиями, важно также признать возможности, которые они открывают в сфере обеспечения безопасности, наращивания оборонного потенциала и обеспечения экономического роста, производительности, устойчивого развития, технологической компетентности и отраслевых инвестиций.

### **Возможности**

Революционные технологии открывают широкие возможности для развития различных отраслей, стимулирования перехода к экологически чистой энергетике, обеспечения устойчивого экономического роста и повышения эффективности и производительности. Кроме того, они способны укреплять безопасность, образование, благосостояние и здравоохранение на глобальном уровне.

ИИ и другие новейшие технологии открывают возможности для укрепления обороноспособности, но в то же время кардинально меняют характер войн будущего, а также средства и методы ведения войны. Технологический прогресс позволяет повысить эффективность сбора и обработки информации, обеспечить более высокий уровень ситуационной осведомленности, ускорить процесс принятия решений, а также принимать более выверенные решения и на более долгосрочную перспективу. Значение дистанционно управляемых и автономных безоператорных систем в современной войне возрастает, и такие системы изменят будущее войны, военные действия и театр военных действий. Все более важное значение будет приобретать прогнозирование развития технологий, интегрирование новейших технологий в оборонные системы и использование неограниченно открывающихся возможностей, поскольку темпы технологического развития в будущем ускорятся. Кроме того, технологическое превосходство может компенсировать численное превосходство.

### **Проблемы**

В то же время важно обеспечить широкое понимание угроз безопасности, потенциала неправомерного использования, проблем с правами человека и факторов взаимозависимости, связанных с развитием таких революционных технологий, как ИИ. По мере их развития они будут порождать новые проблемы, в частности, для оборонного сектора и сектора обеспечения безопасности. Развитие ИИ повышает целенаправленность и эффективность кибератак, деятельности по оказанию информационного воздействия и одного из ее инструментов — дезинформации. Кроме того, ИИ уже используется для того, чтобы влиять на выборы. В этих условиях необходимо уделять повышенное внимание вопросам сохранения конфиденциальной информации.

Нормы международного права, в частности Устав Организации Объединенных Наций, нормы международного права прав человека и международного гуманитарного права, в полной мере распространяются на киберпространство. Уважение и соблюдение рамок ответственного поведения государств в киберпространстве остаются существенно важным условием поддержания международного мира, безопасности и стабильности. Технологический прогресс порождает новые вопросы. Эти вопросы касаются, например, киберсреды, использования ИИ, новых оружейных технологий и разработки критически важных сырьевых ресурсов. Деятельность по оказанию гибридного влияния может включать в себя мероприятия, призванные воспрепятствовать обеспечению подотчетности в соответствии с нормами международного права. Финляндия выступает за то, чтобы в процессе разработки и применения ИИ и составления соответствующих нормативных актов в обязательном порядке учитывались фундаментальные права и права человека, а также сопутствующие риски. Для обеспечения ответственного применения ИИ в военной области в соответствии с нормами международного права важное значение имеет разработка национальных принципов, стандартов и норм, политики и рамочных документов.

Технологический прогресс предоставляет враждебным субъектам новые возможности для осуществления гибридной деятельности по оказанию влияния без пересечения той черты, за которой начинается открытый конфликт. Враждебные кибероперации стали неотъемлемой частью силовой политики и набора инструментов влияния, имеющихся в арсенале государственных субъектов. Кибер-, гибридные и информационные операции проводятся и в обычных условиях, что, со своей стороны, может стирать границы между войной и миром. Несмотря на все более технологичный характер войны, потенциал для ведения обычных боевых действий сохраняет важное значение, особенно в крупномасштабных и долгосрочных конфликтах.

Многие страны сталкиваются с активными действиями по оказанию информационного влияния, в которых также задействуется ИИ. Вредоносное использование информации стало неотъемлемой частью широкомасштабного воздействия, а конкуренция в информационной среде усилилась.

Развитие инфраструктуры и технологический прогресс, а также рост числа пользователей открывают все больше возможностей для ведения враждебных действий в киберпространстве. Многие страны постоянно сталкиваются с проблемами сбора разведанных в информационных сетях, кибершпионажем и кибератаками со стороны враждебных субъектов, которые также стремятся оказать физическое воздействие на объекты критически важной инфраструктуры. Наряду с государственными субъектами все более активную роль в качестве организаторов враждебной деятельности играют политически мотивированные или управляемые государством негосударственные субъекты.

## Франция

[Подлинный текст на французском языке]  
[11 апреля 2025 года]

### I. Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности

#### Возможности, которые стоит изучить

**Помощь в планировании и принятии решений.** Французские вооруженные силы работают над использованием своих баз данных о событиях, связанных с боеприпасами и взрывчатыми веществами, для создания инструментов прогнозирования потенциальных угроз в конкретных областях.

**Оказание поддержки человеку.** С целью улучшить подготовку французских пилотов с использованием анализа данных, полученных в ходе полетов или применения авиационных тренажеров, на основе искусственного интеллекта (ИИ) была разработана система для подготовки летных экипажей. ИИ также может помочь человеку с обработкой больших объемов данных: система «Орей д'ор», например, обрабатывает значительный объем акустических данных, с тем чтобы французские операторы могли фокусироваться на особо важных сигналах.

**Устранение факторов уязвимости в области информационно-коммуникационных технологий.** Технологии ИИ можно использовать для обеспечения кибербезопасности и борьбы с распространением ложной информации. Французские вооруженные силы широко используют системы обнаружения дипфейков.

**Содействие соблюдению международного гуманитарного права и защите людей и имущества.** ИИ может способствовать реализации основополагающих принципов международного гуманитарного права, таких как принципы избирательности, пропорциональности и предосторожности. Помимо этого, ИИ можно использовать для защиты людей, проводя очистку территорий от мин с помощью беспилотных летательных аппаратов, оснащенных датчиками и технологиями ИИ.

**Укрепление контроля над вооружениями.** ИИ можно использовать для обеспечения более эффективного мониторинга и обнаружения тайных пусков и изменений на объектах по производству оружия или проведению испытаний химического и биологического оружия. Наряду с этим благодаря использованию ИИ можно повысить отслеживаемость экспортных поставок вооружений, усилив тем самым контроль за такими поставками.

**Укрепление мер в области предотвращения, а также в области поддержания и упрочения мира.** ИИ позволит повысить уровень адаптивности операций по поддержанию мира и, соответственно, уровень их эффективности. Созданная французскими вооруженными силами система мгновенного перевода «Резистанс» призвана обеспечивать коммуникацию с местным населением в автономном режиме и без доступа к сети, что позволяет бороться с дезинформацией.

#### Риски, которые необходимо снизить

**Риски, связанные с технологией.** Методы обучения сопряжены с различными рисками, порождаемыми предвзятостью: непреднамеренная предвзятость; преднамеренная предвзятость; предвзятость, связанная с реконфигурацией особо чувствительных данных; непрозрачные или труднообъяснимые

результаты. Возникает также проблема существенного роста потребления энергетических ресурсов.

**Усугубление рисков для международной безопасности и стабильности.** При попадании в руки злоумышленников ИИ может усугубить некоторые риски для международной безопасности и стабильности (сценарии эскалации, гонка вооружений, распространение среди негосударственных субъектов, расширение влияния и враждебные действия в киберпространстве), для устранения которых необходимо будет обеспечить соответствующую адаптацию мер по снижению рисков. Наличие риска, связанного с отсутствием возможности возложить ответственность на конкретного субъекта и вызванного зависимостью от технологий, диктует необходимость гарантировать ответственность человека.

## **II. Ключевые принципы и меры для обеспечения «ответственного искусственного интеллекта» на протяжении всего жизненного цикла**

### **Разработка искусственного интеллекта, обеспечивающего уважение международного гуманитарного права**

**Адаптация юридической экспертизы.** Такие экспертизы в полной мере применимы к военному ИИ, но предусматриваемые ими методы должны быть адаптированы к особенностям этой технологии.

**Проведение соответствующих последующих экспертиз.** Эти экспертизы должны проводиться по мере необходимости на различных этапах жизненного цикла оружейной системы. Они должны проводиться тогда, когда устройство модернизируется или когда в него интегрируются новые компоненты, способные существенно изменить последствия его применения.

### **Создание надежного и безопасного искусственного интеллекта**

**Оценка, классификация и сертификация систем.** Эти системы должны проходить процедуры оценки и классификации на должном уровне (с учетом степени важности функций) с проведением анализа рисков на этапе проектирования. Они должны быть связаны с определенными сценариями использования. Эти проверки должны проводиться повторно с той периодичностью, которая необходима с учетом важности соответствующих вопросов.

**Опора на контролируемые и суверенные данные.** Необходимо принять меры по борьбе с рисками утечки данных и создать соответствующие защитные механизмы.

**Внесение корректировок и переобучение систем.** Важно выявлять любые возникающие (во время испытания или эксплуатации) ошибки и давать им характеристику, информировать операторов о необходимости обратной связи и постоянно проверять, отвечает ли система нашим международным обязательствам.

### **Обеспечение надлежащего человеческого контроля над искусственным интеллектом и его включение в ответственную систему подчинения**

**Обеспечение соответствия решений и действий правовым положениям.** Оператор или командующий должен иметь возможность убедиться в том, что предполагаемые результаты согласуются с отданными приказами и правовыми обязательствами.

**Гарантия обеспечения ответственности человека.** Ответственность человека при разработке, внедрении и использовании технологий ИИ — неотъемлемый принцип, требующий формализации иерархии подотчетности для тех, кто отвечает за функции командования, контроля и исполнения.

**Адаптация человеческого контроля.** Анализ и определение характеристик надлежащего человеческого контроля, не ограничивающего возможности системы, использующей технологии ИИ, — сложный вопрос, который должен учитывать различные человеческие, технические и контекстуальные факторы.

**Обучение командующих и личного состава навыкам надлежащего использования этих систем.** Перед использованием необходимо организовать обучение и подготовку, с тем чтобы представители личного состава осознавали преимущества и риски.

### **Разработка устойчивого искусственного интеллекта**

**Защита исследовательской деятельности.** Цели и масштабы исследовательских программ должны быть открытыми и не должны предполагать автоматического введения слишком широких запретов.

**Осмысление этических аспектов исследований.** Во Франции создан постоянный орган — Комитет по этике в оборонном секторе — для рассмотрения этических вопросов, связанных с новыми технологиями в оборонном секторе.

**Разработка низкозатратного искусственного интеллекта.** Выбор в пользу низкозатратных методов предусматривает продумывание использования искусственного интеллекта и повышение устойчивости и рациональности соответствующих систем при одновременном контроле расходов.

## **III. Специальный процесс создания системы глобального регулирования, направленной на практическое применение принципов ответственного искусственного интеллекта**

**Универсальный и инклюзивный процесс.** В обсуждениях должны участвовать все заинтересованные стороны; в частности, крайне важно обеспечить активное участие государств, которые разрабатывают и используют эти системы; с учетом этого в процессе принятия решений необходимо будет учитывать различные позиции и принимать соответствующие правила для обеспечения консенсуса, а также учитывать мнения отраслевых, научных и академических кругов и гражданского общества, с тем чтобы обсуждения не были оторваны от реальности, а подходы были новаторскими. Первый комитет Генеральной Ассамблеи может стать подходящей для этого площадкой.

**Упорядоченная и согласованная архитектура регулирования.** Единый механизм должен позволить упорядочить усилия по повышению эффективности и усилению воздействия полученных результатов. Важно обеспечить взаимодействие с обсуждениями, ведущимися Группой правительственных экспертов по вопросам, касающимся новых технологий в сфере создания смертоносных автономных систем вооружений, которая должна иметь возможность продолжить свою работу после 2026 года в рамках нового мандата.

**Процесс, имеющий оперативную направленность и сосредоточенный на вопросах, касающихся конкретно военного сектора.** Система регулирования должна основываться на своде правовых норм, применимых к вооруженным конфликтам, и, следовательно, в первую очередь — на положениях международного гуманитарного права. Главной задачей любого международного процесса должно быть обеспечение соблюдения существующих правовых норм путем

рассмотрения вопроса о разработке руководящих принципов и способов их реализации государствами (содействие обмену передовым опытом и наращивание международного сотрудничества и помощи с использованием методов, адаптированных с учетом особенностей военного сектора), а также принятие соответствующих мер по укреплению доверия и снижению рисков.

## Германия

[Подлинный текст на английском языке]

[11 апреля 2025 года]

### I. Введение

В последние годы наблюдается беспрецедентное развитие технологий искусственного интеллекта (ИИ), в том числе разработка приложений на основе таких революционных технологий, таких как генеративный ИИ. Необходимо, чтобы государства могли использовать возможности, открывающиеся благодаря этим технологическим достижениям, и не допускать создания препятствий на пути технологического прогресса. В то же время государствам необходимо обеспечивать ответственный подход к разработке и применению приложений ИИ в военной области в полном соответствии с нормами международного права, включая международное гуманитарное право. Для достижения такого равновесия огромное значение имеет международный обмен.

В этих условиях Германия активно участвует в международных процессах, касающихся проблем, которые имеют отношение к ответственному применению ИИ в военной области. В частности, Германия в составе основной группы авторов продвигала резолюции [79/239](#) Генеральной Ассамблеи о применении искусственного интеллекта в военной области и его последствиях для международного мира и безопасности и полностью поддерживает усилия Генерального секретаря по представлению субстантивного доклада с изложением мнений государств-членов о «возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта в военной области».

Германия приветствует эту возможность более подробно изучить мнения государств-членов и других заинтересованных сторон и поделиться элементами своих собственных соображений касательно решения этих важных вопросов.

### II. Принципы и рабочие предположения

Подход Германии к обеспечению ответственного применения ИИ в военной области основывается на изложенных ниже фундаментальных принципах, сформулированных в рамках различных международных форумов и дискуссий.

Германия активно участвовала в разработке принципов ответственного использования, согласованных Организацией Североатлантического договора (НАТО) в 2021 году, и по-прежнему полностью придерживается этих важных стандартов: законность разработки и использования приложений ИИ; ответственность человека за обеспечение подотчетности при разработке и эксплуатации приложений ИИ в военных системах; объяснимость и отслеживаемость приложений ИИ в военной области; надежность, безопасность, защищенность и эффективность систем с ИИ и функциями автономности на протяжении всего их жизненного цикла; и обеспечение управляемости для надлежащего взаимодействия человека с машиной и уменьшения предвзятости.

Кроме того, Германия одобрила итоговые документы двух саммитов по ответственному использованию искусственного интеллекта в военной сфере, состоявшихся в Гааге в 2023 году («Призыв к действию») и в Сеуле в 2024 году («План действий»), а также Политическую декларацию об ответственном применении ИИ и автономных систем в военных целях, инициированную Соединенными Штатами Америки в 2023 году, и активно участвует в осуществлении этой декларации.

Германия участвует также в осуществлении инициативы «Партнерство в сфере ИИ в интересах обороны», в рамках которой страны-единомышленники занимаются поощрением ответственного использования ИИ, продвигают общие интересы и передовой опыт внедрения этики ИИ, создают рамки для содействия сотрудничеству и координируют стратегические сообщения по вопросам политики в сфере ИИ.

В феврале 2025 года Германия одобрила Парижскую декларацию о сохранении человеческого контроля в оружейных системах, использующих ИИ, в которой подчеркивается важность обеспечения человеческого контроля при применении ИИ в военной области.

### **III. Ключевые аспекты, касающиеся применения искусственного интеллекта в федеральных вооруженных силах Германии**

Федеральные вооруженные силы Германии (бундесвер) изучают возможность применения ИИ как для выполнения своих основных задач, так и для достижения превосходства в сферах информации, принятия решений и обеспечения эффективности, а также для оптимизации процессов управления и тылового обеспечения и процессов, касающихся профилактического обслуживания сложных систем. Кроме того, ИИ используется при анализе массовых данных и составлении прогнозов развертывания в помощь специалистам по раннему выявлению гражданских-военных кризисов в различных сферах. ИИ является неотъемлемой частью крупных оборонных проектов, которые также реализуются в европейском контексте, способствуя сохранению и укреплению европейского технологического превосходства. Что касается национальных и технологических разработок в международной оружейной отрасли, то ИИ используется для наращивания потенциала, необходимого для обеспечения национальной и союзнической обороны в будущем. Проработка возможностей применения ИИ, в частности для защиты национальной безопасности и в военных целях, осуществляется в рамках полномочий и обязанностей соответствующих министерств и ведомств. Без ущерба для сказанного выше, аспекты, имеющие отношение к использованию технологий ИИ и приложений ИИ для обеспечения безопасности, включены в стратегию федерального правительства Германии по вопросам ИИ.

Бундесвер предъявляет высочайшие этические требования и устанавливает высочайшие правовые стандарты применительно к использованию ИИ в оружейных системах. В частности, бундесвер руководствуется нормами международного гуманитарного права, применимыми в вооруженных конфликтах, и руководящими принципами Комиссии федерального правительства и НАТО по этике использования данных, в частности упомянутыми выше шести принципами ответственного применения ИИ в военной области на протяжении всего жизненного цикла систем ИИ.

#### IV. Основные соображения

Для поддержания необходимого потенциала обороны и сдерживания Германия по-прежнему твердо намеревается использовать возможности, связанные с применением ИИ в военной области, и убеждена в недопустимости препятствования технологическому прогрессу, особенно учитывая присущий этим технологиям характер, который делает их технологиями двойного назначения.

При этом Германия будет и далее расширять базу знаний путем оценки и устранения рисков, связанных с применением ИИ в военной области, включая риски, связанные с непреднамеренной предвзятостью, например, по гендерному признаку. В этой связи Германия придает большое значение существенно важной роли научных кругов и ценному вкладу научно-исследовательских институтов и аналитических центров, работающих в этой области. В целях содействия проведению соответствующих исследований Германия поддерживает тематические исследовательские организации, в том числе Институт Организации Объединенных Наций по исследованию проблем разоружения (ЮНИДИР), путем внесения финансовых взносов на проведение целевых исследовательских проектов.

Для Германии крайне важно обеспечить всеохватный характер дискуссий, как в географическом плане, так и в плане учета мнений не только государств-членов, но и представителей отраслевых структур, гражданского общества и научных кругов.

Рассматривая возможности и риски, связанные с использованием ИИ в системах вооружений, Германия придает особое значение концепции человеческого контроля и считает наличие эффективной системы сохранения человеческого контроля необходимым условием для обеспечения соответствия всех систем вооружений нормам международного гуманитарного права. Это подразумевает наличие не только технического контроля, но и определенной степени рассудительности. Разработанная в Германии концепция системы человеческого контроля включает набор технологически осуществимых шагов и действий, которые устанавливают четкие границы допустимого для алгоритма системы. Центральное место в этих границах занимают нормы международного права, в частности международного гуманитарного права. В процессе реального применения ИИ на поле боя огромное значение имеет контекст. Германия считает, что концепция системы человеческого контроля обеспечивает надлежащий способ должного учета этого контекста.

Особое внимание необходимо уделять вопросам применения ИИ в увязке с ядерным оружием, то есть той области, в которой научные и политические дебаты все еще находятся на ранних стадиях. Возможное применение ИИ в системах командования и управления ядерным оружием может иметь серьезные последствия для стратегической стабильности или ядерной эскалации. В то же время ИИ может открыть новые возможности для сдерживания распространения и применения оружия массового уничтожения. Германия стремилась внести свой вклад в эти дебаты, организовав конференцию по вопросам искусственного интеллекта и оружия массового уничтожения в рамках получившей широкое признание серии конференций по теме «Овладение технологиями — переосмысление контроля над вооружениями», которая прошла в Берлине 28 июня 2024 года.

Конвенция о запрещении разработки, производства и накопления запасов бактериологического (биологического) и токсинного оружия и об их уничтожении и Конвенция о запрещении разработки, производства, накопления и применения химического оружия и о его уничтожении запрещают целые категории



оружия массового уничтожения. Такие виды применения, как большие языковые модели (генеративные), могут способствовать распространению знаний двойного назначения, которые могли бы неправомерно использоваться для разработки, производства или применения биологического и химического оружия. Сочетание таких видов применения ИИ как AlphaFold и синтетическая биология может обеспечить злоумышленникам возможность проектировать новые белки, которые благодаря изменениям в последовательности ДНК могут ускользать от обнаружения. ИИ может использоваться для анализа больших массивов облачных данных, например данных о геноме человека, и приносить значительную пользу в разработке индивидуальных методов лечения, но может применяться и для создания биологического оружия, нацеленного на конкретные этнические группы.

Поэтому в тесном сотрудничестве со своими международными партнерами Германия продолжит выявлять возможные направления действий для оценки влияния приложений ИИ на разработку и производство запрещенного оружия и вводить возможные правила. В то же время Германия будет использовать преимущества ИИ для целей проверки, проведения биоэкспертизы и уменьшения рисков.

## V. Приверженность Германии международным процессам

Германия с самого начала активно участвует в процессе «Ответственное применение искусственного интеллекта в военной области» и будет делать это и далее. Германия вошла в основную группу авторов резолюции [79/239](#) Генеральной Ассамблеи о применении искусственного интеллекта в военной области и его последствиях для международного мира и безопасности. Германия высоко оценивает межрегиональный и многосторонний подход, принятый в рамках указанной важной инициативы, и с нетерпением ждет ее продолжения в Испании в сентябре 2025 года.

Действуя строго последовательно, Германия внесла свой вклад в инициированную Соединенными Штатами Политическую декларацию об ответственном применении ИИ и автономных систем в военных целях, в частности стала председателем рабочей группы по надзору (совместно с Австрией).

Кроме того, Германия активно участвует в работе Партнерства по ИИ в целях обороны и Экспертной сети ЮНИДИР по управлению искусственным интеллектом в военной области.

Германия поддерживает посла Роберта ин ден Боша, Председателя базирующейся в Женеве Группы правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений, и продолжает активно участвовать в соответствующем процессе, в том числе координируя позиции нескольких государств-членов в рамках так называемой двухуровневой группы. В тесном сотрудничестве со своими международными партнерами Германия будет и далее работать над тем, чтобы эта группа своевременно выполнила свой мандат, желательно к концу 2025 года.

В контексте НАТО Германия признает потенциал ИИ для дальнейшего развития вооруженных сил и наращивания оборонного потенциала альянса, а также отмечает проблемы, которые создаст применение ИИ для оперативной совместимости вооруженных сил стран, входящих в альянс. Необходимо всесторонне учитывать многонациональные разработки в области ИИ и аспекты стандартизации ИИ в рамках НАТО, Европейского союза и стран — партнеров Германии, чтобы обеспечить оперативную совместимость бундесвера в качестве военной силы в контексте международных операций. Поэтому Германия приветствует

тот факт, что страны НАТО согласовали принципы ответственного применения в контексте стратегии НАТО по вопросам ИИ.

## VI. Дальнейшие шаги

Поскольку новейшие революционные технологии будут и далее развиваться и определять картину нашего мира, Германия считает необходимым обеспечить всеобъемлющую международную координацию в сфере ответственной разработки и применения ИИ в военной области. Протекающие международные процессы обеспечивают прекрасную основу для рассмотрения соответствующих значимых аспектов и обеспечения учета мнений самых различных заинтересованных сторон. Германия будет и далее активно поддерживать эти усилия в целях претворения в жизнь и расширения поддержки таких политических обязательств по ответственному применению ИИ в военной области, как Политическая декларация, принятая по инициативе Соединенных Штатов, и процесс «Ответственное применение искусственного интеллекта в военной области». Германия с нетерпением ожидает того момента, когда она сможет изучить результаты доклада Генерального секретаря об искусственном интеллекте в военной области. Германия будет и впредь активно участвовать в процессе, касающемся смертоносных автономных систем вооружений, который протекает в рамках Группы правительственных экспертов в Женеве.

## Греция

[Подлинный текст на английском языке]  
[10 апреля 2025 года]

Внедрение искусственного интеллекта (ИИ) в оборонном секторе фундаментальным образом повлияло на средства и методы ведения военных операций. Применение ИИ в военной области дает значительные оперативные преимущества, в том числе повышает скорость принятия решений, позволяет совершенствовать выявление и прогнозирование угроз, обеспечивает ситуационную осведомленность и оценку в режиме реального времени, оптимизирует распределение и планирование ресурсов и тыловое обеспечение, расширяет возможности личного состава при выполнении сложных задач и позволяет эффективно обрабатывать обширные массивы разведывательных данных.

Тем не менее, несмотря на указанные преимущества, существенно важно признать, что технический прогресс также порождает сложные, многоплановые проблемы, которые требуют тщательного изучения во избежание нанесения ущерба миру, безопасности и стабильности на региональном и глобальном уровнях.

В этой связи Греция считает, что одна из ключевых областей, вызывающих беспокойство, заключается в применении военных систем с возможностью машинного обучения, которое порождает ряд проблем, включая транспарентность и объяснимость, поскольку сложные модели могут работать по принципу «черного ящика» с неопределенными процессами принятия решений, особенно в условиях постоянно меняющейся обстановки на поле боя.

Кроме того, потенциальное применение генеративного ИИ в боевой технике создает дополнительный уровень сложности и неопределенности, поскольку такие системы способны автономно генерировать новые решения и адаптироваться к меняющимся условиям на поле боя, постоянно анализируя новые данные и обучаясь на их основе, что для Греции является главным поводом для беспокойства. Для устранения этих проблем крайне важно установить

четкие оперативные границы и ограничения на применение таких систем во избежание непреднамеренного поведения.

С учетом сказанного выше одна из самых тревожных проблем, связанных с применением ИИ в военных целях, заключается в его внедрении в системы командования и управления и поддержки принятия решений о применении ядерного оружия. Перспектива передачи системам с поддержкой ИИ полномочий на принятие решений, связанных с ядерным сдерживанием, или даже инициирования соответствующих протоколов применения требует тщательного рассмотрения в целях обеспечения человеческого надзора и участия человека в принятии таких решений, а также установления соответствующих гарантий кибербезопасности во избежание непреднамеренной эскалации.

Не меньшую озабоченность в нынешней сложной геополитической обстановке вызывают усилия государств по сохранению военного превосходства, то есть усилия, которые могли бы подстегнуть гонку вооружений, характеризующуюся отсутствием транспарентности и взаимной подозрительностью. Это соперничество может усугубить геополитическую нестабильность и создать серьезные проблемы для глобальной безопасности, поскольку баланс сил нарушается, а технологический разрыв между развитыми и развивающимися государствами все более расширяется.

Кроме того, все более активная разработка и развертывание вооруженными силами средств с поддержкой ИИ может привести к понижению порога вооруженного конфликта. Ускорение темпов принятия решений и все более широкое применение безоператорных систем на поле боя повышают риск непреднамеренной эскалации, поскольку человеческий фактор на поле боя все чаще заменяется безоператорными системами.

В этом контексте нужно надлежащим образом рассмотреть еще один аспект, который заключается в распространении и перенаправлении средств с поддержкой ИИ государствам, игнорирующим основанный на правилах международный порядок, и негосударственным субъектам, включая террористические организации. По мере повышения доступности технологий ИИ возникает значительный риск их приобретения и использования такими субъектами в дестабилизирующих целях, что еще больше подорвет международную безопасность.

Кроме того, военные применения ИИ порождают риски и проблемы, связанные с проведением психологических операций и дезинформацией, поскольку они позволяют массово генерировать дезинформацию, дипфейки и поддельные данные с целью обмана общественности и дестабилизации институциональных структур. Использование учетных записей, управляемых автоматизированными системами (ботами), и алгоритмов целевой пропаганды позволяет усиливать психологические операции, в частности влиять на общественное мнение и избирательные процессы и создавать социальную напряженность, в том числе подрывая доверие населения к операциям по поддержанию мира посредством кампаний дезинформации. Кроме того, возникают опасения по поводу социальных предубеждений, например в отношении гендерных аспектов, возраста, расовой принадлежности и инвалидности, поэтому существенно важно проводить оценку рисков и принимать меры по их уменьшению во избежание непреднамеренной предвзятости и дискриминации в алгоритмах.

Более того, приложения с использованием ИИ по обеспечению кибербезопасности могут использоваться как для защиты объектов критически важной инфраструктуры, так и в злонамеренных целях, например, для совершения кибератак и перехвата данных. Гибридные угрозы, сочетающие традиционные военные операции с наступательными разведывательными операциями, требуют

повышения бдительности и усиления координации действий государственных и международных субъектов, чтобы избежать эскалации и сохранить региональный и международный мир и безопасность.

В свете сказанного выше Греция решительно поддерживает международные усилия по обеспечению ответственного применения ИИ в военной области, поскольку, несмотря на описанные выше проблемы, оно может способствовать обеспечению соблюдения норм международного гуманитарного права и защиты гражданского населения путем повышения точности целеуказаний, совершенствования обзорной разведки и оптимизации гуманитарной помощи.

В этом духе 4 апреля 2025 года Греция совместно с Францией и Республикой Корея, а также при ценной поддержке Армении, Италии и Королевства Нидерландов организовала заседание Совета Безопасности по формуле Аррии на тему «Использование безопасных, инклюзивных, заслуживающих доверия технологий искусственного интеллекта для поддержания международного мира и безопасности». Это заседание позволило получить ценные сведения о том, каким образом Организация Объединенных Наций может способствовать поддержанию международного мира и безопасности, особенно посредством регулирования, обеспечения нераспространения и недопущения перенаправления средств ИИ в военную область, укрепления верховенства права, демократических ценностей, социальной сплоченности и экономического развития.

Кроме того, на международной арене Греция поддержала совместные заявления о мерах по ответственному развитию и применению ИИ в военной области, принятые на двух саммитах по ответственному использованию искусственного интеллекта в военной сфере, состоявшихся в Гааге (15 и 16 февраля 2023 года) и Сеуле (9 и 10 сентября 2024 года). Греция поддержала также принятую по инициативе Соединенных Штатов Америки Политическую декларацию об ответственном применении ИИ и функции автономности в военной области и Парижскую декларацию о сохранении человеческого контроля в оружейных системах, использующих ИИ.

Более того, Греция создала консультативный комитет высокого уровня<sup>1</sup> по вопросам ИИ для разработки комплексной национальной стратегии в области ИИ, а также необходимые структуры в Министерстве национальной обороны для решения технологических, правовых, этических и политических проблем, возникающих в связи с применением ИИ и функции автономности в военной области.

И наконец, в стремлении внести конструктивный вклад в международный диалог по вопросу об ответственном применении ИИ в военной области Греция организует международную конференцию по теме «Вооруженные конфликты и управление кризисами в эпоху ИИ», которая пройдет в Афинах 22 и 23 мая 2025 года.

<sup>1</sup> В подготовленном Комитетом знаменательном исследовании под названием “A blueprint for Greece’s AI transformation” («План преобразования Греции в сфере ИИ») содержатся руководящие принципы и описываются передовые проекты для развития искусственного интеллекта в Греции, в которых задачами первостепенной важности являются защита и укрепление демократии, смягчение последствий изменения климата и адаптация к ним, а также поддержка безопасности.

## Индия

[Подлинный текст на английском языке]  
[1 апреля 2025 года]

Искусственный интеллект (ИИ) представляет собой преобразующую технологию, которая оказывает существенное влияние на все аспекты человеческой жизни. Он развивается с беспрецедентным размахом и беспрецедентными темпами, оперативно внедряется и используется в различных сферах. ИИ может оказать преобразующее воздействие в плане снижения уровня бедности и улучшения жизни людей. Это особенно актуально для развивающихся стран, таких как Индия.

Необходимы коллективные глобальные усилия для создания системы управления ИИ и выработки соответствующих стандартов в поддержку наших общих ценностей, для устранения рисков и укрепления доверия. Системы управления ИИ и соответствующие стандарты должны обеспечивать учет глубокой трансграничной взаимозависимости, поощрять инновации, применяться на благо всего мира, а также содействовать обеспечению доступа и справедливости, чтобы блага использования ИИ были доступны всем, особенно странам глобального Юга. Индия решительно выступает за открытое обсуждение вопросов инноваций и управления.

Дискуссии по вопросам военного ИИ должны быть привязаны к военной реальности, для которой характерно оперативное внедрение ИИ в военные доктрины и операции. Продолжающиеся конфликты по всему миру демонстрируют как риски, так и возможности, связанные со все более широким внедрением этих технологий.

Разработка, внедрение и применение ИИ в военной области сопряжены с проблемами этики, права и безопасности. Не умаляя значения этих проблем, Индия поддерживает ту точку зрения, что ИИ может способствовать соблюдению норм международного гуманитарного права.

Индия поддерживает коллективные глобальные усилия по надлежащему регулированию разработки, внедрения и применения ИИ в военной области. Эти усилия должны быть направлены на решение правовых и этических проблем, а также на выявление и уменьшение рисков, связанных с применением ИИ в военной области.

Любые коллективные усилия по надлежащему регулированию применения ИИ в военной области должны быть направлены на применение и использование, а не на саму технологию и ее составные компоненты. Необходимо избегать стигматизации технологий. Доступ к технологиям для целей развития не должен ограничиваться.

ИИ следует применять в военной области на законных основаниях в соответствии с неотъемлемым правом на индивидуальную или коллективную самооборону, предусмотренным нормами международного права. Нормы международного гуманитарного права по-прежнему в полной мере распространяются на применение ИИ в военной области. Кардинальные принципы международного гуманитарного права, а именно избирательность, пропорциональность и предосторожность, применимы ко всем средствам и методам ведения войны в прошлом, настоящем и будущем.

Человеческое суждение и надзор за применением ИИ в военной области имеют существенно важное значение для уменьшения рисков и обеспечения соблюдения норм международного гуманитарного права.

В любых коллективных усилиях или соответствующих нормативных актах, касающихся применения ИИ в военной области, следует учитывать существующие правовые обязательства и уважать национальную юрисдикцию и компетенцию, а также учитывать соответствующий национальный потенциал.

Индия привержена ответственному применению ИИ в военной области.

Индия разрабатывает систему оценки надежности ИИ в оборонном секторе для решения сложных задач, связанных с современными технологиями ИИ. В основе этой системы лежат пять ключевых принципов: а) надежность и отказоустойчивость; b) физическая и техническая безопасность; c) транспарентность; d) непредвзятость; e) приватность. Эти принципы закладывают основу для дальнейших дискуссий по вопросу о надлежащем регулировании разработки, внедрения и применения ИИ в военной области.

## Индонезия

[Подлинный текст на английском языке]

[11 апреля 2025 года]

Индонезия приветствует обсуждение возможностей и проблем для международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, с уделением особого внимания областям помимо смертоносного автономного оружия, как это предусмотрено в пунктах 7 и 8 резолюции 79/239 Генеральной Ассамблеи.

Поскольку в военной области применяются самые различные системы и приложения ИИ, для организации всеобъемлющего многостороннего обсуждения этой темы в Организации Объединенных Наций необходимо обеспечить рассмотрение вопросов, касающихся не только средств кинетического поражения (таких как смертоносные автономные системы вооружений), но и некинетических средств, которые могут оказывать непосредственное воздействие (например, автономные боевые киберсистемы, адаптивные системы постановки противорадиолокационных помех или средства радиоэлектронной борьбы) или вспомогательное воздействие в военных целях (например, тыловое обеспечение, медицинская эвакуация или тактическая разведка). Такое обсуждение должно охватывать и другие средства, которые могут оказывать непосредственное влияние на стратегическое равновесие, например в плане совершенствования дистанционного зондирования (спутниковый контроль или поиск подводных лодок), ведения разведки или планирования военных действий.

Индонезия по-прежнему твердо привержена делу поддержания международного мира и безопасности, как это закреплено в преамбуле к Конституции Индонезии. В силу этой приверженности Индонезия считает, что применение ИИ в военной области должно регулироваться таким образом, чтобы способствовать укреплению мира, безопасности и достижению целей в области устойчивого развития. ИИ должен стать силой на службе мира и безопасности, а не элементом, подрывающим безопасность, разжигающим конфликты или подпитывающим стратегическое соперничество.

Хотя сам по себе ИИ не является оружием, Индонезия признает, что он служит как мультипликатором силы, так и фактором усиления угроз, способным давать значительные преимущества и создавать серьезные риски для международного мира и безопасности. Применение ИИ в военной области порождает различные этические, правовые, моральные и технические вопросы, которые необходимо тщательно рассмотреть и обсудить с точки зрения соблюдения норм

международного права, включая международное гуманитарное право и международное право прав человека.

С одной стороны, считается, что ИИ обладает широким потенциалом. Он может повышать эффективность обработки данных, повышать степень эффективности, точности и безошибочности операций и потенциально способствовать соблюдению норм международного гуманитарного права, например путем содействия оценке соразмерности и мер предосторожности ради уменьшения вреда гражданскому населению. Кроме того, ИИ может способствовать наращиванию потенциала разведки, наблюдения и рекогносцировки, поддерживать тыловое обеспечение и планирование, а также улучшать управление личным составом.

С другой стороны, ИИ порождает целый ряд факторов риска и последствий, включая вероятность разжигания гонки вооружений, распространение среди негосударственных субъектов, вероятность преступного и безответственного применения, усиление диспропорций в военной мощи за счет обеспечения технологического превосходства, а также усиление нестабильности, повышение вероятности просчетов и эскалации и правовой неопределенности. К числу факторов технического риска относятся также киберуязвимость, сбои в работе систем, необъективность данных, ошибочное опознавание целей и другие неопределенности оперативного уровня.

Индонезия особо обеспокоена экзистенциальными рисками, возникающими в результате возможного внедрения ИИ в системы командования, управления и связи ядерных сил. Индонезия подтверждает свою принципиальную позицию, в соответствии с которой применение и угроза применения ядерного оружия нарушают нормы международного права, и заявляет о необходимости принятия срочных и решительных мер для сохранения и усиления норм, запрещающих ядерное оружие. Внедрение ИИ в ядерные оружейные системы усугубляет экзистенциальные риски применения ядерного оружия, будь то преднамеренное, непреднамеренное или случайное применение, и повышает уровень ядерной опасности. Это представляет собой угрозу безопасности всех стран. Индонезия настоятельно призывает все государства, обладающие ядерным оружием, пересмотреть свою зависимость от ядерного оружия и подтвердить нашу коллективную приверженность делу построения мира, свободного от ядерного оружия. В период до полной ликвидации ядерного оружия государства, обладающие ядерным оружием, должны сохранять значимый человеческий контроль, ответственность и подотчетность в отношении ядерного оружия и систем его доставки в контексте развития ИИ.

С учетом этих соображений Индонезия настоятельно призывает применять осторожный подход к решению проблем, связанных с применением ИИ в военной области. Индонезия подчеркивает необходимость регулирования разработки, применения и использования ИИ в военной области в интересах заедствования его преимуществ и уменьшения связанных с ним рисков. Такое регулирование должно служить делу обеспечения коллективного мира, безопасности и процветания всех стран. В этой связи Индонезия предлагает перечисленные ниже ключевые соображения.

Во-первых, Индонезия заявляет, что нормы международного права необходимо соблюдать на протяжении всего жизненного цикла технологий ИИ. К их числу относятся Устав Организации Объединенных Наций, нормы международного гуманитарного права, международного права прав человека, а также договоры о разоружении и нераспространении. Государства должны проводить соответствующую правовую экспертизу на всех этапах — от закупки до оценки. Государства должны гарантировать ответственный подход к своей деятельности

по разработке и применению ИИ в военной области, в том числе законность применения ИИ в процессе ведения войны или военных действий. Важно подчеркнуть, что в отсутствие законов, регулирующих применение ИИ в военной области, такое применение должно регулироваться законами человечности и велениями общественной совести.

Помимо международного права, правовые рамки регулирования вопросов применения ИИ в военной области должны дополняться этическими соображениями. В процессе разработки и применения средств ИИ необходимо обеспечивать соблюдение таких принципов, как отслеживаемость, подотчетность, ответственность, объяснимость, гуманность, транспарентность, равенство и справедливость.

Во-вторых, Индонезия подчеркивает существенно важную роль человеческого фактора в обеспечении подотчетности и ответственности на всех уровнях, будь то государственный, корпоративный или индивидуальный, при проектировании, разработке, внедрении и применении ИИ в военной области.

Разработка, применение и использование ИИ в военной области неизменно должны носить человекоцентричный характер и служить интересам человечества. Необходимо сохранять эффективный и значимый человеческий контроль, особенно при принятии решений, связанных с применением силы, и усиливать его посредством подготовки кадров. При принятии критически важных решений необходимо обеспечивать участие в принятии решений, вмешательство, надзор и контроль со стороны человека. Кроме того, хотя «значимый человеческий контроль» все чаще используется в процессе регулирования применения ИИ в военной области, Индонезия считает, что эта концепция пока не отвечает правовым, моральным, техническим и нормативным аспектам, связанным с таким применением. Необходимо договориться о том, что на практике подразумевает понятие «значимый» применительно к человеческому контролю.

Хотя управление ИИ будет в первую очередь регулировать поведение государственных структур, оно должно охватывать и гражданских участников, в частности технологические компании, причастные к применению ИИ в военной области. Государства должны обеспечивать соблюдение частным сектором норм международного права и этических стандартов, поддерживая при этом развитие инновационной экосистемы ИИ. Научные сотрудники и компании несут ответственность за обеспечение надежности, физической и технической безопасности, подотчетности и подконтрольности человеку создаваемых ими технологий ИИ. Они должны отвечать также за мониторинг, информирование и устранение рисков, присущих их продукции.

В-третьих, Индонезия подчеркивает настоятельную необходимость принятия многосторонней, инклюзивной и всеобъемлющей нормативно-правовой системы управления. Она должна отражать интересы всех государств, независимо от уровня развития ИИ в этих государствах. В процессе разработки правил и норм, регулирующих применение ИИ в военной области, необходимо в равной мере учитывать голоса всех государств, чтобы обеспечить справедливое представительство и повысить уровень глобального доверия.

Участие широкого круга заинтересованных сторон имеет решающее значение, учитывая многоплановость этических, правовых и технических аспектов ИИ. Кроме того, прежде чем применять системы ИИ в военной области, необходимо обеспечить их соответствие нормам международного права, гуманитарного права, правам человека и обязательствам в области разоружения путем взаимодействия специалистов по различным дисциплинам и представителей различных культур.



В-четвертых, крайне важно сохранять осведомленность и стимулировать содержательное обсуждение рисков, проблем и последствий, связанных с разработкой, внедрением и применением ИИ в военной области, как в технологических, так и не в технологических сферах. Индонезия подчеркивает важность постоянной оценки более широких последствий военного применения ИИ для международного мира и безопасности, особенно в контексте нераспространения и разоружения. Для понимания этих последствий, которые до сих пор недостаточно изучены, требуются более комплексные исследования.

Выявление рисков, связанных с разработкой, внедрением и применением ИИ в военной области, будет способствовать прогнозированию на основе фактических данных, оценке рисков и, в конечном счете, разработке мер по их уменьшению.

Крайне важно также обеспечивать более глубокое понимание рисков, связанных с применением ИИ в военной области, и повышать уровень осведомленности о них. В этой связи следует добиваться повышения уровня транспарентности, в том числе путем обмена информацией о национальной политике и стратегии, особенно для выявления, оценки и уменьшения рисков, совместного использования средств ИИ в военной области, сообразно обстоятельствам, в целях повышения подотчетности и укрепления доверия, и обмена накопленным опытом и примерами передовой практики между странами, отраслевыми структурами и секторами.

В-пятых, управление ИИ не должно препятствовать технологическому развитию или ограничивать доступ к ИИ для развивающихся стран. Системы регулирования не должны навязывать условий или создавать препятствия, ограничивающие равноправный доступ. Необходимо выработать сбалансированный подход, в котором учитывались бы такие риски, как распространение, и который одновременно обеспечивал бы доступность ИИ для государств с ограниченными ресурсами.

И наконец, в процессе управления ИИ необходимо уделять особое внимание преодолению отставания в сфере цифровых технологий и ИИ. Развивающиеся страны сталкиваются с серьезными проблемами не только в плане доступа к потенциалу ИИ, но и в плане способности эффективно управлять этими технологиями. Если не устранить этот пробел, то это подорвет усилия по глобальному управлению, поскольку многие государства по-прежнему не готовы решать сложные и трансграничные задачи, которые ставит перед ними ИИ.

Индонезия подчеркивает настоятельную необходимость преодоления значительного отставания в сфере цифровых технологий и ИИ на уровне стран и в самих странах, особенно в сфере доступа к финансовым, людским и техническим ресурсам. Это отставание может привести к углублению глобального неравенства и усилению конфликтного потенциала.

Мир и безопасность в качестве глобального общественного блага требуют международного сотрудничества между всеми странами, как развитыми, так и развивающимися, для решения общих проблем и извлечения коллективных выгод, в том числе связанных с разработкой, применением и использованием ИИ в военной области. В этой связи Индонезия призывает наладить расширенное и сбалансированное международное сотрудничество и оказание помощи в деле наращивания глобального потенциала ИИ и развития механизмов управления. Такое сотрудничество должно осуществляться на справедливой и взаимосогласованной основе с учетом конкретных потребностей и условий развивающихся стран. Это включает, в частности, инициативы по наращиванию потенциала, образованию, передаче технологий, обучению на протяжении всей жизни,

подготовке технических кадров, проведению совместных исследований и обмену знаниями.

Такое сотрудничество должно носить многоуровневый характер и охватывать не только государства и международные организации, но и внутригосударственные секторы. Необходимо поощрять налаживание государственно-частных партнерств, чтобы способствовать выработке ответственного подхода к инновациям и повышать уровень осведомленности представителей отраслевых структур о последствиях, которые разрабатываемые ими технологии могут иметь для международного мира и безопасности.

Международное сотрудничество имеет решающее значение не только для преодоления отставания в сфере цифровых технологий и ИИ, но и для создания условий, благоприятствующих укреплению доверия в отношениях между государствами. Это поможет способствовать понижению уровня геополитической разобщенности и соперничества в сфере ИИ. Международное сотрудничество должно основываться на принципах равенства, доверия, взаимной выгоды, уважения суверенитета и солидарности, чтобы открывать путь к полноценному взаимодействию, включая передачу технологий и обмен знаниями.

Кроме того, Индонезия признает ценность укрепления механизмов регионального сотрудничества с учетом местной и региональной специфики. Такие механизмы могут служить основополагающими составными элементами для выработки более широкого глобального консенсуса, одновременно обеспечивая форум для проведения более детальных обсуждений с учетом существующего контекста.

## **Иран (Исламская Республика)**

[Подлинный текст на английском языке]  
[12 марта 2025 года]

В соответствии с просьбой, обращенной к Генеральному секретарю в пункте 7 резолюции 79/239, в котором Генеральная Ассамблея запросила мнения государств-членов о возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений, Исламская Республика Иран настоящим представляет свои мнения.

Искусственный интеллект становится одним из главных движущих факторов перемен в современном мире, в значительной мере определяя развитие военной промышленности на ближайшее будущее и тем самым налагая существенный отпечаток на картину международного мира и безопасности. Государственные и негосударственные субъекты активно продвигают свои разнообразные программы в области ИИ, которые нельзя оставлять без регулирования. Учитывая ведущую роль негосударственных субъектов и необходимость соблюдения баланса между регулятивными и инновационными процедурами и тенденциями, крайне важно, чтобы полномочия по регулированию оставались суверенной прерогативой государств-членов.

С точки зрения существа этого вопроса Исламская Республика Иран, как и в случае с другими технологиями, используемыми в киберпространстве и в космосе, поддерживает исключительно мирное применение ИИ, учитывая, что в соответствующих обстоятельствах военные структуры также могут пользоваться дивидендами ИИ в мирных целях.

Принимая во внимание различия в уровне развития разных стран, крайне важно не допустить, чтобы отставание в сфере цифровых технологий перешло в сферу ИИ. Инклюзивность всех процедур регулирования, связанных с ИИ, может быть гарантирована только на основе консенсуса в рамках Организации Объединенных Наций. Такой подход обеспечивает суверенитет государств-членов, способствует созданию условий для справедливого развития ИИ для всех и обеспечивает инновационную гибкость для процветания индустрии ИИ. Центральная роль Организации Объединенных Наций в вопросах регулирования ИИ препятствует применению эксклюзивных национальных подходов к этому вопросу. В этом жизненно важном вопросе должны царить инклюзивность и консенсусный подход.

Несмотря на текущие обсуждения вопросов ИИ на различных международных форумах, наше понимание этого вопроса и его последствий для международного мира и безопасности остается неполным. Было бы преждевременно утверждать, что нормы международного права, гуманитарного права и международного права прав человека в полной мере применимы к ИИ. В свете огромных масштабов этого нового и быстро развивающегося явления могут потребоваться меры по адаптации и развитию соответствующей международно-правовой базы.

Что касается международных усилий по регулированию, то в качестве предпочтительного направления действий Исламская Республика Иран поддерживает идею принятия государствами-членами юридически обязывающих соглашений, а не нормотворческих или политических документов.

В силу своей принципиальной позиции по вопросам разоружения Исламская Республика Иран отвергает любые политически мотивированные, дискриминационные или условные подходы, равно как и двойные стандарты. Поэтому терминология, используемая Генеральной Ассамблеей, должна отражать стремление к единству и консенсусу. В этом смысле такие понятия, как «ответственное применение», являются слишком абстрактными, чтобы регулировать сферу, которая определяется конкретно и точно. Столь абстрактное понятие может неверно истолковываться и создавать предпосылки для применения политизированного подхода. Исламская Республика Иран выражает свое решительное несогласие с использованием такой субъективной терминологии. Она выступает за то, чтобы в любом будущем документе заменить термин «ответственное применение» на «мирное применение», и вносит такое предложение.

## Израиль

[Подлинный текст на английском языке]  
[10 апреля 2025 года]

Израиль приветствует принятие резолюции [79/239](#) и в соответствии с пунктом 7 этой резолюции имеет честь представить свои мнения для включения в доклад Генерального секретаря, который должен быть представлен Ассамблее на ее восьмидесятой сессии для дальнейшего обсуждения государствами-членами.

Израиль придерживается того мнения, что концепция искусственного интеллекта (ИИ) в настоящее время допускает целый ряд возможных толкований, которые могут быть уточнены с течением времени.

Очевидно, что ИИ в военной области применяется все шире и чаще, чем когда-либо прежде. Израиль проголосовал за указанную резолюцию Генеральной Ассамблеи и призывает государства и все заинтересованные стороны принять участие в обсуждении, сохраняя при этом профессиональный и неполитизированный подход и учитывая законные интересы всех государств, в том числе соображения безопасности, гуманитарного характера, экономики и развития.

Мы считаем, что для проведения серьезной и ответственной дискуссии по вопросу о применении ИИ в военной области, которая в конечном итоге сможет принести значимый результат, необходимо принять прагматичный, сбалансированный и поэтапный подход.

Поскольку технологии открывают широкие возможности практически во всех областях, в том числе и в военной, мы приветствуем изучение благ, которые могут принести эти разработки, и способов их практической реализации, а также потенциальных рисков и способов их уменьшения. Израиль считает, что такие новые технологии, как технологии ИИ, могут также способствовать соблюдению действующих норм международного гуманитарного права. В интересах реализации таких потенциальных возможностей необходимо не допускать негативной стигматизации этих технологий.

Израиль по-прежнему принимает конструктивное участие в глобальной дискуссии по вопросу о применении ИИ в военных целях. Недавно он поддержал Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях, инициированную Соединенными Штатами. Мы с нетерпением ожидаем возможности принять участие в предстоящих встречах в рамках этой инициативы и продолжить поддержку ответственного применения ИИ и автономных систем в военных целях.

В последние годы в рамках этой декларации, а также в других контекстах государства рассматривают возможность выработки руководящих принципов по вопросам развития и применения искусственного интеллекта в военной области как на национальном, так и на международном уровне. Некоторые из наиболее базовых и общепринятых принципов таких руководящих принципов, которые могут также иметь отношение к обсуждениям в контексте резолюции [79/239](#), как представляется, заключаются в следующем:

- военное применение ИИ должно соответствовать применимым нормам международного права.
- Оно должно быть ответственным и способствовать укреплению международной безопасности.
- Государства должны обеспечивать подотчетность в отношении использования средств ИИ в соответствии с применимыми нормами международного права, в том числе путем сохранения человеческого контроля в рамках ответственной цепи командования и управления.

К числу практических мер, которые государствам следует принимать для претворения в жизнь этих принципов, относятся следующие:

- государствам следует предпринимать соответствующие шаги, такие как проведение правовых обзоров, для обеспечения того, чтобы применение их военных средств ИИ согласовывалось с их соответствующими обязательствами по международному праву, в частности международному гуманитарному праву.

- Государствам следует принимать соответствующие меры для обеспечения ответственного подхода к разработке, развертыванию и применению военных средств ИИ. Такие меры должны применяться на соответствующих этапах на протяжении всего жизненного цикла военных средств ИИ.
- Соответствующий персонал должен проявлять должную осмотрительность при разработке, развертывании и применении военных средств ИИ, включая системы вооружений, в которых используются такие средства.
- Старшие должностные лица должны осуществлять эффективный и надлежащий надзор за разработкой и развертыванием военных средств ИИ, применение которых может иметь серьезные последствия, включая, в частности, системы вооружений с использованием таких средств.
- Государствам следует поддерживать соответствующие усилия по обеспечению ответственного и законного применения военных средств ИИ и последовательно обсуждать с другими государствами вопросы, касающиеся способов развертывания и применения военных средств ИИ в этом качестве.

Израиль считает целесообразным проведение всесторонних многосторонних дискуссий по вопросу о применении ИИ в военной области и его последствиях для международной безопасности, чтобы изыскать правильный баланс между военной необходимостью и гуманитарными соображениями.

## Италия

[Подлинный текст на английском языке]

[11 апреля 2025 года]

### Председательство Италии в Группе семи

В период председательства Италии в Группе семи в 2024 году вопросам искусственного интеллекта (ИИ) уделялось главное внимание в ходе политических и технических дискуссий. В ходе состоявшейся в Апулии встречи на высшем уровне было признано влияние ИИ на военную область и необходимость создания системы регулирования для обеспечения ответственного подхода к его разработке и применению.

С 18 по 20 октября 2024 года в Неаполе прошла первая встреча стран Группы семи на уровне министров обороны. В ходе этой встречи министры обороны стран Группы семи вновь подтвердили свое твердое намерение согласованным и конкретным образом решать проблемы безопасности в нынешний период истории, который характеризуется высоким уровнем нестабильности. Кроме того, они подчеркнули необходимость выработки более кооперативного подхода к проведению научно-исследовательских и опытно-конструкторских работ в оборонной сфере, в том числе в плане обмена опытом и знаниями и их использования, а также создания безопасной среды для предотвращения доступа со злонамеренными целями, ради сохранения конкурентных преимуществ, в том числе в сфере новых и революционных технологий.

И наконец, Группа директоров по нераспространению стран — членов Группы семи опубликовала заявление, в котором признала глубокое влияние таких новых революционных технологий, как ИИ, на контроль над вооружениями, нераспространение и разоружение, а также на военные операции в будущем.

## **I. Процесс, касающийся ответственного применения искусственного интеллекта в военной области**

Италия высоко оценивает процесс, касающийся ответственного применения искусственного интеллекта в военной области, который был инициирован Нидерландами и Республикой Корея в 2023 году с целью создания платформы для обсуждения ключевых возможностей, проблем и рисков, связанных с военным применением ИИ. На втором саммите по ответственному использованию искусственного интеллекта в военной сфере, который прошел в Сеуле в 2024 году, Италия одобрила документ под названием «План действий», в котором излагаются ключевые принципы ответственного управления ИИ, включая важность соблюдения норм международного права, обеспечения ответственности и подотчетности операторов, безотказности и надежности систем ИИ, а также надлежащего участия человека в разработке, внедрении и применении ИИ в военной области.

Государства, которые одобрили этот документ, подчеркивают необходимость предотвращения использования технологий ИИ для содействия распространению оружия массового уничтожения и важность ненанесения ущерба усилиям по контролю над вооружениями, разоружению и нераспространению. Кроме того, в целях выработки общего понимания технологии ИИ и ее применения в военной области в «Плане действий» к государствам обращен призыв взять на себя обязательства по дальнейшему обсуждению и разработке эффективных процедур проведения правовой экспертизы, а также принятия мер по укреплению доверия и соответствующих мер по уменьшению рисков. В этом контексте для достижения прогресса в обсуждениях решающее значение имеют обмен информацией и передовым опытом, а также активное участие других заинтересованных сторон.

## **II. Политическая декларация об ответственном применении искусственного интеллекта и автономных систем в военных целях**

Италия высоко оценивает *Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях*. Подтвердив, что военное применение ИИ может и должно быть этичным и ответственным и способствовать укреплению международной безопасности, государства, поддержавшие Декларацию, признали, что при разработке, внедрении и применении военных средств ИИ необходимо принять комплекс мер. В частности, государства обязались: сводить к минимуму непреднамеренную предвзятость в военных средствах ИИ, обеспечивать надлежащую и тщательную проверку их физической и технической безопасности и эффективности, а также внедрять соответствующие меры предосторожности для выявления и предотвращения непредвиденных последствий и эффективного реагирования в соответствующих случаях. Кроме того, важно определить систему ответственного командования и управления и использовать военные средства ИИ в соответствии с международными обязательствами.

## **III. Пакт во имя будущего**

В сентябре 2024 года мировые лидеры приняли «Пакт во имя будущего», подтвердив свои глобальные обязательства и предоставив государствам возможность действовать сообразно новым и новейшим проблемам и перспективам. В действии 27 государствам рекомендуется пользоваться возможностями, возникающими в связи с новыми и новейшими технологиями, включая ИИ, и в то же время устранять потенциальные риски, связанные с их неправомерным использованием. В частности, государства-члены будут и далее оценивать такие риски,

присущие военным приложениям ИИ, и потенциальные возможности на протяжении всего их жизненного цикла на основе консультаций с соответствующими заинтересованными сторонами.

#### **IV. Парижская декларация о сохранении человеческого контроля в оружейных системах, использующих искусственный интеллект**

Кроме того, Италия недавно поддержала *Парижскую декларацию о сохранении человеческого контроля в оружейных системах*, использующих искусственный интеллект, которая была принята в рамках саммита по искусственному интеллекту, прошедшего в Париже 6–11 февраля 2025 года. Государства, поддержавшие эту декларацию, подчеркнули, что ответственность и подотчетность никогда нельзя перекладывать на машины, и обязались применять человекоцентричный подход к разработке, внедрению и применению приложений ИИ в военной области. Они также обязались обеспечивать, чтобы внедрение ИИ в военной сфере полностью соответствовало нормам международного права и международного гуманитарного права, а также поощрять научно-исследовательские и опытно-конструкторские работы и новаторские разработки с использованием технологий ИИ.

#### **V. Группа правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений**

Стремительное развитие искусственного интеллекта и машинного обучения также имеет существенные последствия для использования функции автономности в системах вооружений. Италия считает, что Конвенция о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие, в качестве документа, в котором отражен совокупный дипломатический, правовой и военный опыт представителей правительств, международных организаций и специализированных учреждений, безусловно, обеспечивает наиболее подходящий форум для решения текущих и возникающих вопросов, связанных с разработкой и применением систем вооружений. Италия активно участвует в обсуждениях, проводимых в рамках Группы правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений, которая была создана под эгидой этой конвенции, и твердо намерена продвигать дискуссии по вопросу о разработке элементов будущего документа в соответствии с мандатом, согласованным на Сессии Высоких договаривающихся сторон Конвенции в 2023 году.

Италия считает, что этот документ должен содержать четкие запреты и правила, чтобы в конечном итоге его можно было принять в качестве дополнительного протокола к Конвенции. Согласно такому подходу, смертоносные автономные системы вооружений, которые не могут разрабатываться и применяться в соответствии с нормами международного гуманитарного права, должны быть *ipso facto* запрещены. С другой стороны, системы с автономным принятием решений в части критически важных функций, которые могут разрабатываться и применяться в полном соответствии с нормами международного гуманитарного права, подлежат нормативному регулированию. По мнению Италии, человеческий фактор по существу имеет решающее значение для всего жизненного цикла смертоносных автономных систем вооружений, а именно для их проектирования, разработки, производства, развертывания и применения. Необходимо также сохранить надлежащий уровень участия человека в принятии решений и осуществлении контроля, чтобы обеспечить ответственность и подотчетность в соответствии с нормами международного гуманитарного права.

## Япония

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

В резолюции 79/239 Генеральная Ассамблея просила Генерального секретаря запросить мнения государств-членов и государств-наблюдателей о возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений, и представить Генеральной Ассамблее на ее восьмидесятой сессии для дальнейшего обсуждения государствами субстантивный доклад, содержащий обобщение этих мнений и каталог существующих и новых нормативных предложений, а также приложение с изложением этих мнений. Настоящим Япония представляет свои мнения по этому вопросу, чтобы внести вклад в подготовку доклада и в дальнейшее обсуждение этой темы.

### I. Общие замечания

Япония привержена поддержанию и укреплению свободного и открытого международного порядка, основанного на верховенстве права, чтобы все люди могли наслаждаться благами мира, стабильности и процветания, а также поощрению дипломатии ради построения безопасного и прочного мира, в котором обеспечивается защита человеческого достоинства. Руководствуясь этими целями, Япония активно участвует в усилиях по укреплению международного мира и безопасности, а также контролю над вооружениями и разоружением.

Япония считает, что вопрос о применении ИИ в военной области должен рассматриваться на комплексной основе, при обеспечении достаточного понимания соответствующих рисков и преимуществ, а также с учетом как гуманитарных соображений, так и аспектов безопасности. Представляется целесообразным углубить понимание в вопросах применения ИИ в военной области и способствовать приложению реалистичных и практических усилий для его ответственного применения, чтобы максимально расширить его преимущества и одновременно уменьшить риски.

Что касается применения ИИ в военной области, то Япония поддерживает мнение о том, что, во-первых, действующие нормы международного права применяются к регулируемым им вопросам, которые возникают на протяжении всего жизненного цикла ИИ; во-вторых, средства ИИ должны применяться ответственно; и, в-третьих, ответственность за их применение и последствия по-прежнему лежит на человеке. Япония подчеркивает также необходимость повышения степени транспарентности в качестве важной меры укрепления доверия в интересах получения максимальных преимуществ при одновременном уменьшении рисков.

### II. Мнения и подход Японии в том, что касается возможностей и проблем в деле обеспечения международного мира и безопасности в связи с применением искусственного интеллекта в военной области

#### Возможности

##### *Мнения*

Стремительный научно-технический прогресс, включая развитие ИИ, кардинально меняет парадигму безопасности. Страны стремятся разработать передовые технологии, которые могли бы кардинальным образом изменить характер войны и, соответственно, радикально изменить ситуацию, причем на практике



стало крайне сложно проводить различия между технологиями для гражданского использования и технологиями для целей обеспечения безопасности. ИИ обладает колоссальным потенциалом для преобразования всех аспектов военного дела, включая военные операции, командование и управление, ведение разведки, наблюдения и рекогносцировки, подготовку личного состава, управление информацией и тыловое обеспечение. С учетом многообразия видов применения ИИ в военной области он может способствовать, в частности, повышению точности, кучности и эффективности, повышению уровня ситуационной осведомленности и понимания ситуации, повышению оперативности анализа информации, уменьшению количества ошибок, совершаемых человеком, и экономии трудозатрат. Надлежащее применение ИИ может способствовать совершенствованию защиты гражданского населения в условиях конфликта и постконфликтному миростроительству.

*Принятый Японией подход к использованию возможностей*

В процессе применения ИИ в военной области необходимо учитывать, насколько эффективно такое применение способствует преодолению выявленных человеком проблем, не забывая при этом о функциях и ограничениях ИИ. Применение ИИ не должно являться самоцелью, и его не следует рассматривать в отрыве от его функций и ограничений. Поэтому государствам следует обеспечивать, чтобы военные средства ИИ имели четко определенное применение, а также разрабатывались и проектировались для выполнения этих установленных функций. В этой связи важно способствовать формированию общего международного понимания ИИ, его функций и ограничений в военной области, а также общего понимания потенциального применения ИИ в военной области. Что касается применения ИИ военными ведомствами, то в июле 2024 года Министерство обороны Японии опубликовало основы политики Министерства обороны по содействию применению ИИ, в котором изложило свои текущие соображения о функциях и ограничениях ИИ в военной области, а также о тех сферах применения ИИ, которым оно уделяет первостепенное внимание. В этих основах политики Министерство обороны, учитывая текущие возможности и ограничения ИИ, определило следующие семь сфер применения ИИ, которым оно уделяет особое внимание:

- Выбор и опознавание целей
- Сбор и анализ разведывательных данных
- Командование и управление
- Тыловое обеспечение
- Безоператорные системы
- Кибербезопасность
- Повышение эффективности административной работы

В этих основах политики также отмечается необходимость исходить из того, что ИИ применяется для оказания человеку поддержки в принятии решений и что участие человека в процессе применения ИИ имеет существенно важное значение.

## Проблемы

### Мнения

Применение ИИ в военной области чревато тем, что оно может перейти в неправомерную сферу или приобрести злонамеренный характер, а также рисками эскалации и понижения порога конфликта, которые могут возникнуть в результате предвзятости, непреднамеренных последствий и других факторов. В этой связи Япония подчеркивает необходимость не допускать использования ИИ для содействия распространению оружия массового уничтожения государствами и негосударственными субъектами и подчеркивает, что ИИ должен поддерживать усилия в области разоружения, контроля над вооружениями и нераспространения, а не препятствовать им.

### *Принятый Японией подход к решению проблем*

С учетом таких рисков, как предвзятость, неправомерное и злонамеренное использование, Министерство обороны Японии будет заниматься вопросами уменьшения рисков, связанных с ИИ, на базе концепций человекоцентричности ИИ, обеспечения технической безопасности, справедливости, неприкосновенности частной жизни, физической безопасности, транспарентности и подотчетности, сформулированных в Руководстве по ИИ для коммерческих структур в Японии, которое было опубликовано в апреле 2024 года, обращая при этом внимание на дискуссии, протекающие в международном сообществе и проводимые с оборонными ведомствами других стран.

Кроме того, Япония внимательно следит за возможными последствиями применения таких новых технологий, как ИИ для ядерного разоружения и нераспространения. В этой связи Япония приветствует тот факт, что в ходе Конференции 2022 года участников Договора о нераспространении ядерного оружия по рассмотрению действия Договора Соединенные Штаты, Соединенное Королевство и Франция взяли на себя обязательство сохранять контроль и участие человека во всех действиях, имеющих критически важное значение для информационного обеспечения и осуществления суверенных решений, касающихся применения ядерного оружия, и призывает другие государства, обладающие ядерным оружием, последовать этому примеру. Кроме того, Международная группа видных деятелей за мир без ядерного оружия в своей рекомендации обзорной конференции 2026 года подчеркнула необходимость совместной работы по решению проблем и использованию возможностей, связанных с новейшими технологиями.

## III. Мнения о будущих дискуссиях и международном сотрудничестве

Для управления ИИ в военной области необходим гибкий, сбалансированный и реалистичный подход, который позволит не отставать от стремительных темпов развития и совершенствования технологий. Япония подчеркивает, что усилия по обеспечению ответственного применения ИИ в военной области могут предприниматься параллельно с усилиями по проведению научно-исследовательских и опытно-конструкторских работ, экспериментальных и новаторских работ в сфере технологий ИИ и не препятствовать им.

Следует отметить, что проведение дискуссий, направленных на поиск проблем, связанных с конкретными технологиями ИИ, может затормозить технологическое развитие и инновационные процессы в гражданском секторе и оказать сдерживающее воздействие. Кроме того, вопросы применения ИИ в военной области должны обсуждаться на всесторонней основе с привлечением заинтересованных сторон и проведением обмена мнениями между ними.

С учетом сказанного выше Япония решительно поддерживает итоговые документы саммитов по ответственному использованию искусственного интеллекта в военной сфере и Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях и выражает надежду на то, что к этим инициативам присоединится больше государств.

Следует отметить, что в вопросе о смертоносных автономных системах вооружений Япония решительно выступает за продолжение обсуждений в рамках Конвенции о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие, и ожидает, что обсуждения вопросов применения ИИ в военной области дополнят и укрепят дискуссии в Группе правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений, которая была создана в рамках Конвенции.

Япония признает, что обеспечение транспарентности в сфере применения ИИ в военной области имеет важное значение в качестве меры укрепления доверия, ведущей к уменьшению рисков, а также к обеспечению эффективного взаимодействия и сотрудничества между странами. Япония признает также важность наращивания потенциала для содействия принятию ответственного подхода к разработке, внедрению и применению ИИ в военной области и преисполнена решимости укреплять международное сотрудничество в сфере наращивания потенциала, нацеленное на сокращение пробелов в знаниях о таком подходе. В этой связи представляются полезными такие методы, как обмен передовым опытом и информацией об извлеченных уроках, и Япония намеревается использовать возможности для обмена мнениями с другими странами.

И наконец, что касается применения ИИ в военной области, Япония будет и далее активно и конструктивно участвовать в международных дискуссиях с целью выработки общего понимания в международном сообществе путем проведения сбалансированных обсуждений, учитывающих гуманитарные соображения и перспективы безопасности.

## Литва

[Подлинный текст на английском языке]  
[9 апреля 2025 года]

Литва выражает признательность за эту возможность представить материалы для доклада Генерального секретаря согласно резолюции [79/239](#) Генеральной Ассамблеи. Литва с удовлетворением поддержала эту резолюцию, которую Генеральная Ассамблея приняла 24 декабря 2024 года.

Литва отмечает, что развитие и применение искусственного интеллекта (ИИ) в военной области создает как возможности, так и проблемы для международного мира и безопасности. Литва придает большое значение разработке норм и принципов ответственного использования, которые позволят государствам использовать преимущества и смягчить потенциальные риски применения ИИ в военной области. Литва твердо убеждена в том, что все ответственные государства заинтересованы в обеспечении ответственного применения ИИ в военной области. Литва считает, что для обеспечения учета последствий военного применения ИИ необходимы глобальные действия и многосторонний подход с участием государственного и частного секторов, гражданского общества и научных кругов.

Литва всецело поддерживает Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях, к которой Литва присоединилась 13 ноября 2023 года. Эта политическая декларация содержит принципы, не имеющие юридически обязывающего характера, и примеры передовой практики, призванные обеспечить ответственное и законное применение ИИ в военном контексте. В Политической декларации учитываются такие меры, как проведение правовой экспертизы, обеспечение надлежащего надзора, сведение к минимуму непреднамеренной предвзятости и разработка четких и надлежащим образом сформулированных сценариев применения военных средств ИИ. Литва настоятельно рекомендует большему числу государств подписать эту политическую декларацию.

Кроме того, Литва присоединилась к стратегии Организации Североатлантического договора по вопросам ИИ, которая была принята в 2021 году и пересмотрена в 2024 году. В этой стратегии сформулированы шесть принципов ответственного применения ИИ в военной области, а именно: законность, ответственность и подотчетность, объяснимость и отслеживаемость, надежность, управляемость и уменьшение предвзятости. Эти не имеющие юридически обязывающего характера принципы, которые решительно поддерживает Литва, должны применяться на протяжении всего жизненного цикла приложений ИИ.

И наконец, Литва с удовольствием делится своими мнениями относительно возможностей и проблем в сфере международной безопасности, создаваемых применением ИИ в военной области. Литва считает, что военные средства ИИ могут и должны применяться ответственно, в первую очередь ради укрепления национальной безопасности государства, а также для содействия соблюдению норм международного права, в том числе международного гуманитарного права, и для выполнения различных обязательств государства по защите гражданского населения. Помимо усиления защиты гражданских лиц в вооруженном конфликте ответственное применение ИИ открывает возможности для совершенствования процесса принятия решений, тылового обеспечения, планирования и принятия других мер повышения эффективности.

Что касается потенциальных рисков применения ИИ в военной области, то Литва хотела бы выделить, в частности, такие проблемы, как кибербезопасность, непреднамеренная предвзятость военных средств ИИ и непредусмотренное поведение систем с поддержкой ИИ. Литва считает, что наилучшим способом устранения таких потенциальных рисков является внедрение принципов ответственного применения, а также наращивание потенциала и надлежащая подготовка личного состава по использованию приложений ИИ и систем с поддержкой ИИ. Литва подчеркивает, что для получения доступа к преимуществам ИИ в военной области и использования ИИ в качестве критически важного компонента оборонного потенциала, государствам следует избегать введения неоправданных чрезмерных ограничений, препятствующих инновациям в сфере ИИ, особенно если безответственные государства отказываются принимать любые такие ограничения в отношении военного применения ИИ.

## Мексика

[Подлинный текст на испанском языке]  
[10 апреля 2025 года]

### **Искусственный интеллект, автономные системы вооружений и проблемы их регулирования, с которыми сталкивается весь мир**

Мексика представляет настоящий документ в соответствии с резолюцией 79/239 Генеральной Ассамблеи, озаглавленной «Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности».

Мексика признает, что применение искусственного интеллекта (ИИ) в военной области может приносить пользу. Вместе с тем ИИ также создает серьезные проблемы для международного мира и безопасности, которые требуют срочного и скоординированного внимания со стороны международного сообщества.

Мексика высоко ценит проводимые в рамках Организации Объединенных Наций многосторонние обмены мнениями, способствующие формированию общего понимания возникающих рисков и общей ответственности, например «Первый диалог по вопросам военного ИИ: возможности, риски и международный мир и безопасность», координируемый Управлением Организации Объединенных Наций по вопросам разоружения. Мы считаем, что внедрение ИИ в военные приложения порождает фундаментальные проблемы для международного мира и безопасности, включая непреднамеренную эскалацию конфликта, стратегическую неопределенность и все более широкое применение силы автономными системами.

Мексика считает первостепенной задачей укрепление международного сотрудничества, содействие обеспечению транспарентности, обмен передовым опытом, наращивание потенциала в целях формирования культуры соблюдения и уважения норм международного права, а также достижение прогресса в создании нормативной базы, обеспечивающей регулирование разработки и внедрения ИИ в военной области этическими, правовыми и гуманитарными принципами во избежание углубления диспропорций или нарушения международной стабильности.

### **Международный мир и безопасность**

Мексика считает необходимым принятие мер по предотвращению распространения и неправомерного применения таких технологий, в том числе негосударственными субъектами или вне четких правовых рамок.

Международный мир и безопасность нельзя ставить в зависимость от внедрения новых и новейших технологий в военной области. В этом процессе необходимо учитывать перспективы развития человеческого потенциала и расширения социальных прав и возможностей, особенно в интересах развивающихся стран. Поэтому при изучении таких технологий следует делать упор на их использование в мирных целях и в целях урегулирования споров, а не на повышение эффективности военной машины.

Повышение уровня сложности цифровых угроз и перспектива использования государствами новых технологий для совершения атак на другие государства, а также сложность обеспечения надежности и точности автономных систем в военных условиях, уязвимость систем ИИ на протяжении их жизненного цикла, алгоритмическая предвзятость, отравление данных и использование

генеративных моделей в злонамеренных целях подчеркивают необходимость принятия упреждающих мер по уменьшению рисков.

Научно-технический прогресс, особенно в сфере ИИ, автономных систем, цифровых и квантовых технологий, опережает нынешние возможности нормативно-правовой базы по управлению такими рисками. Поэтому Мексика вновь подчеркивает необходимость разработки всеобъемлющих рамок регулирования, поощрения международного сотрудничества и многостороннего диалога, а также уделения первостепенного внимания вопросам обеспечения транспарентности, подотчетности и реального контроля со стороны человека на протяжении всего жизненного цикла этих технологий, включая тщательное тестирование и обеспечение этических гарантий в процессе их внедрения.

В отсутствие четких международно-правовых рамок и необходимого многостороннего консенсуса использование термина «ответственный» в данном контексте не должно рассматриваться как молчаливое одобрение или согласие на использование или развитие автономных военных систем с поддержкой ИИ. Принцип ответственности обязательно должен быть связан с законностью и подотчетностью.

В соответствии с такой логикой Мексика считает существенно важным создать механизмы управления и регулирования, понижающие вероятность использования ИИ и других революционных технологий во враждебных целях, признавая, что риски не ограничиваются лишь их оперативным развертыванием, а возникают уже на начальных этапах их проектирования и разработки.

### **Оперативная обстановка**

С учетом того, что применение ИИ в военной области может осуществляться в различной оперативной обстановке, Мексика отмечает, что эта технология может оказывать различное воздействие.

В условиях вооруженного конфликта крайне важно обеспечить, чтобы любые технологии, поддерживаемые ИИ, применялись в соответствии с нормами международного гуманитарного права, в частности принципами избирательности, соразмерности, предосторожности и гуманности.

В рамках операций по поддержанию мира и ликвидации последствий стихийных бедствий ИИ может внести позитивный вклад в координацию материально-технического обеспечения, прогнозирование рисков и оказание помощи пострадавшему населению при условии строгого соблюдения прав человека.

Что касается безопасности границ, то Мексика признает, что ИИ может усилить возможности контроля, однако необходимо подчеркнуть важность обеспечения уважения достоинства всех людей, избегая принятия автоматических решений, закрепляющих дискриминационную практику.

### **Смертоносные автономные системы вооружений**

Мексика считает, что основополагающей частью этой дискуссии являются смертоносные автономные системы вооружений, которые представляют собой особую угрозу для международного мира и безопасности. В этой связи мы выступаем за то, чтобы не фрагментировать многосторонние дискуссии относительно внедрения новых технологий в военной области, и считаем, что вопрос о смертоносных автономных системах вооружений должен быть неотъемлемой частью этих дискуссий.

Мексика считает, что международное сообщество должно в безотлагательном порядке установить четкие запреты и правила в отношении применения смертоносных автономных систем вооружений, поскольку они не согласуются с нормами международного гуманитарного права и порождают риски с точки зрения этики, права и безопасности.

Мексика стала одним из инициаторов и авторов резолюций [78/241](#) и [79/62](#) в Генеральной Ассамблее о смертоносных автономных системах вооружений, преследуя цель укрепить легитимное многостороннее пространство для решения этих проблем.

Мексика поддерживает призыв Генерального секретаря и Международного комитета Красного Креста начать к 2026 году переговоры относительно юридически обязывающего документа, устанавливающего необходимые запреты и правила в отношении смертоносных автономных систем вооружений, как это отражено в Новой повестке дня для мира.

Мексика подтвердила свою политическую приверженность рассмотрению этой темы тем, что она приняла участие в Конференции в Сан-Хосе (2023 год), присоединилась к Беленскому коммюнике, приняла активное участие в конференции по теме «Человечество на перепутье: автономные системы вооружений и проблема регулирования» (Вена, 2024 год) и одобрила ее заключительный доклад.

Мексика считает, что смертоносные автономные системы вооружений порождают многочисленные риски, в частности:

- исключают человеческое суждение при принятии критических решений о применении силы;
- подменяют важнейшую функцию оценки обстановки в военных операциях;
- ослабляют механизмы обеспечения подотчетности и присвоения ответственности.

Ответственность за применение силы ни при каких обстоятельствах нельзя перекладывать на машину. Решения о развертывании, активировании или деактивировании систем вооружений во всех случаях должны приниматься людьми, которые должны нести за это юридическую ответственность.

Мексика вновь заявляет, что применение любых военных технологий, включая технологии, основанные на ИИ, должно осуществляться с соблюдением международных обязательств, вытекающих из:

- Устава Организации Объединенных Наций;
- норм международного гуманитарного права;
- норм международного права прав человека;
- норм международного уголовного права;
- норм права о международной ответственности.

В этой связи Мексика считает необходимым запретить те системы вооружений, которые основаны на технологии:

- неспособной проводить различия между военными и гражданскими целями;
- неспособной применять принцип соразмерности при оценке сопутствующего ущерба;

- не имеющей механизмов отмены в случае, если атака сочтена нецелесообразной;
- причиняющей излишние страдания или чрезмерный вред здоровью людей.

Мексика считает необходимым срочно приступить к переговорам относительно юридически обязывающего документа, который устанавливал бы конкретные запреты и правила в отношении смертоносных автономных систем вооружений, обеспечивал значимый человеческий контроль в критически важных видах применения, а также предусматривал эффективные механизмы внедрения, мониторинга и подотчетности.

### Преимущества и риски

Что касается конкретных видов использования ИИ, то Мексика отмечает как преимущества, так и риски в следующих областях.

- Командование и управление: при определенных условиях ИИ мог бы повысить эффективность принятия оперативных решений, но они должны оставаться под значимым контролем человека, особенно когда речь идет о применении силы. Известно, что ИИ способен обрабатывать и анализировать большие объемы данных и информации, значительно превосходящие возможности человека, что позволяет быстрее, проще и эффективнее прогнозировать будущие тенденции и принимать стратегические решения в режиме реального времени.
- Кибероперации: применение ИИ открывает ценные возможности для прогнозирования инцидентов в киберпространстве и реагирования на них, но при этом усиливает риски эскалации напряженности, включая автоматическое применение мер воздействия без надлежащего надзора.
- Управление информацией и материально-техническое снабжение: обработка больших массивов данных с помощью ИИ может способствовать принятию решений в режиме реального времени, но этот процесс должен осуществляться в соответствии с протоколами, обеспечивающими этическое, подотчетное и ответственное использование этой технологии.

Несмотря на сказанное выше, Мексика особо отмечает технологические риски, связанные с внедрением ИИ в военной области, поскольку имеются сведения о том, что технические сбои или непредвиденные ошибки могут привести к эскалации конфликта.

### Нидерланды (Королевство)

[Подлинный текст на английском языке]  
[7 апреля 2025 года]

Королевство Нидерландов приветствует возможность представить свои мнения по вопросу о проблемах и возможностях в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, как это предусмотрено резолюцией 79/239, принятой Генеральной Ассамблеей 24 декабря 2024 года.

Нидерланды отмечают потенциал военного применения ИИ и решительно выступают за ответственный подход к разработке, внедрению и применению ИИ в военной области. Принципиальная позиция Нидерландов заключается в том, что применение ИИ в военной области должно соответствовать нормам международного права, включая Устав Организации Объединенных Наций, нормы международного гуманитарного права и международного права прав человека.



15 и 16 февраля 2023 года в Нидерландах прошел первый саммит по ответственному использованию искусственного интеллекта в военной сфере. С тех пор процесс, касающийся ответственного применения искусственного интеллекта в военной области, стал платформой для многостороннего партнерства в целях обсуждения ключевых возможностей и проблем, связанных с военным применением ИИ, с участием представителей правительств, научных учреждений, аналитических центров, отраслевых структур и организаций гражданского общества. Обсуждения проводятся раз в год на глобальном уровне, а в течение года — в ходе региональных мероприятий, посвященных вопросам ответственного применения искусственного интеллекта в военной области, которые уже состоялись в Сингапуре, Кении, Турции, Чили и Нидерландах.

На саммите 2023 года Нидерланды и еще 57 стран согласовали совместный призыв о необходимости ответственного подхода к разработке, внедрению и применению ИИ в военной области. В 2024 году Нидерланды одобрили План действий, который был согласован в ходе саммита по ответственному использованию искусственного интеллекта в военной сфере, организованного Республикой Корея совместно с Нидерландами. Кроме того, Нидерланды поддержали Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях.

В ходе саммита 2023 года Нидерланды создали Глобальную комиссию по вопросам ответственного применения искусственного интеллекта в военной области, возложив на нее задачу разработки краткосрочных и долгосрочных рекомендаций для правительств и более широкого круга заинтересованных сторон. Нидерланды ожидают публикации доклада этой комиссии со стратегическими указаниями, который ожидается в сентябре 2025 года.

Ниже кратко излагается позиция Нидерландов и ключевые вопросы, требующие дальнейшего рассмотрения.

### **Возможности для международного мира и безопасности**

С военной точки зрения основными преимуществами ИИ являются скорость и масштаб. Технология ИИ позволяет гораздо быстрее обрабатывать и анализировать данные. Кроме того, системы поддержки прогнозирования боевой обстановки и принятия решений с использованием ИИ помогают командирам определять курс действий. Это позволяет совершенствовать стратегический анализ и повышает способность оперативно и эффективно реагировать на угрозы.

Нидерланды считают, что ИИ может также способствовать обеспечению международного мира и безопасности путем совершенствования информационного анализа, усиления защиты гражданского населения и уменьшения рисков в ходе проведения боевых операций:

- системы поддержки анализа и принятия решений с использованием ИИ улучшают информационное обеспечение командиров касательно оперативной обстановки и долгосрочного стратегического прогноза. Это помогает лучше понять динамику перемещений гражданского населения в зонах конфликта, проблемы климатической безопасности, гендерное насилие и поведенческие модели террористических организаций. Такая информация, в свою очередь, может использоваться для совершенствования управления рисками и деятельности по урегулированию конфликтов, способствуя тем самым укреплению международного мира и безопасности.

- Нидерланды считают целесообразным применение ИИ в военной области для улучшения связи между силами обороны, а также между силами обороны и другими субъектами, например персоналом по оказанию гуманитарной помощи, надзорными организациями и местными органами власти. Обмен данными может осуществляться между большим количеством пользователей, что позволяет создавать единые источники достоверной информации с помощью «умных» датчиков, работающих в защищенной сетевой среде. Кроме того, агенты ИИ могут использоваться для ускорения обмена данными. Улучшение связи за счет увеличения скорости обмена данными способствует обеспечению международного мира и безопасности, повышая эффективность связи, обмена информацией и международного сотрудничества, например применительно к системам раннего оповещения о кризисах и кризисного регулирования.
- Нидерланды придают большое значение потенциалу ИИ для защиты гражданского населения. ИИ способен распознавать закономерности и отклонения в больших объемах данных, что может обеспечивать более полное представление о ситуации в гражданской среде. Такое более глубокое понимание может уменьшить риск ошибочной идентификации, сопутствующего ущерба и жертв среди гражданского населения. В более широком смысле ИИ может способствовать повышению качества работы по выявлению возможных угроз гражданскому населению и гражданским объектам, что позволит вооруженным силам оперативно и адекватно реагировать на них. Кроме того, ИИ может способствовать оптимизации усилий по оказанию гуманитарной помощи, например, в сфере обеспечения продовольствия, жилья и медицинского обслуживания в зонах конфликта. И наконец, ИИ может способствовать расследованию потерь среди гражданского населения путем сбора и анализа данных и доказательств, чтобы определить причину потерь и обеспечить привлечение виновных к ответственности.
- ИИ позволяет уменьшать риски для личного состава, находящегося на передовой, поскольку автономные системы под управлением ИИ могут заменить человека при ведении определенных действий в труднодоступных или опасных районах. В качестве примеров можно привести подводное наблюдение и поддержку поисково-спасательных операций в экстремальных погодных условиях. Помимо этого, ИИ может способствовать уменьшению расходов на оказание медицинской помощи и проведение реабилитации путем уменьшения воздействия на личный состав в условиях повышенного риска.

### **Проблемы для международного мира и безопасности**

Нидерланды считают, что применение ИИ в военной области порождает целый ряд рисков для международного мира и безопасности:

- Нидерланды обеспокоены тем, что ИИ может использоваться для расширения, совершенствования и автоматизации кибератак и манипулирования информацией, поскольку и то и другое подрывает международный мир и безопасность. Развитие генеративного ИИ упрощает манипулирование информацией и проведение автоматизированных кибератак. Его применение в военной области нарушает оперативную связь и усложняет процесс принятия решений. В долгосрочной перспективе широкое распространение дезинформации и проведение автоматизированных кибератак могли бы подорвать доверие к военным линиям связи. Кроме того, это могло бы повлиять на уровень доверия между государствами, потенциально разрушая тем

самым напряженные отношения, особенно между странами, которые уже находятся на грани потенциального конфликта.

- Риски, связанные с применением ИИ в военной области, могут привести к появлению систем, применение которых потенциально противоречило бы нормам международного права. Такие недостатки могут являться следствием неспособности учитывать существующую обстановку и имеющиеся данные или правильно понимать военный жаргон и, в свою очередь, могут привести к чрезмерному упрощению процесса принятия военных решений или, например, игнорированию конкретной оперативной обстановки. Кроме того, если то или иное приложение поведет себя непредсказуемым образом, выдаст дискриминационные результаты, основанные на нерелевантных характеристиках, или предложит незаконный порядок действий, то потенциально это может привести к нарушению международно-правовых обязательств государства. В связи с широким распространением ИИ последствия смещения автоматизации, необъективности данных и человеческих решений, основанных на предвзятости систем ИИ, способны создать значительные проблемы в плане присвоения ответственности, а также обеспечения подотчетности и принятия надлежащих мер по исправлению ситуации. Важно отметить, что от приложений ИИ нельзя ожидать, что они будут рассуждать или действовать как люди.
- Риск эскалации в результате действий ИИ создает потенциальные риски для международного мира и безопасности. Поскольку благодаря повышению быстродействия и расширению возможностей применение ИИ ускоряет цикл «наблюдения, оценки обстановки, принятия решения и нанесения удара», в результате расхождений между военными намерениями и анализом, проводимым системами с поддержкой ИИ, могут возникать недоразумения. Поэтому ИИ может непреднамеренно способствовать эскалации. Поскольку системы ИИ способны обнаруживать цели быстрее и захватывать их в большем количестве по сравнению с человеком, применение таких систем может также повысить интенсивность и смертоносность конфликтов.
- Как следствие, создание надежных систем обороны становится все более сложной задачей. То, с какой скоростью появляются новые приложения ИИ, затрудняет решение стратегических и тактических задач по эффективному противодействию таким системам и защите от них в условиях ведения боевых действий. Это конкретное последствие все более широкого применения систем ИИ может потенциально способствовать ведению наступательных действий и таким образом негативно повлиять на международный мир и безопасность.
- Еще одной проблемой становится дестабилизация по мере того, как террористические организации, организованные преступные сообщества и другие негосударственные субъекты получают доступ к военным средствам ИИ. В этой связи Нидерланды обеспокоены тем, что ИИ может сделать производство химического, биологического, радиологического или ядерного оружия более доступным для таких субъектов.

В свете стремительного развития технологий ИИ Нидерланды признают, что в настоящее время невозможно предвидеть все проблемы и возможности в сфере обеспечения международного мира и безопасности. Некоторые из них являются совершенно новыми, а другие уже существуют, но могут получить дополнительное развитие в результате применения ИИ. Для обеспечения ответственного применения ИИ в военной области всеми государствами необходим постоянный международный диалог по этому вопросу.

## **Ответственное применение искусственного интеллекта в военной области**

Чтобы обеспечить ответственное применение ИИ в военной области, необходимо сохранять участие человека в принятии решений и контроль со стороны человека, соответствующие контексту. Ответственность и подотчетность всегда должны сохраняться за человеком. В этой связи важно отметить ряд моментов, которые излагаются ниже.

*Расширение контроля со стороны человека не гарантирует более высокий уровень ответственности искусственного интеллекта*

Нидерланды считают, что не существует универсального решения для обеспечения достаточного участия человека в принятии решений и обеспечении контроля в приложениях ИИ. В зависимости от ряда факторов, человеческое суждение и контроль варьируются от непосредственного управления человеком до более высоких уровней автоматизации и автономности. Поэтому в каждом конкретном случае необходимо определять степень человеческого суждения и контроля в приложениях и системах, управляемых ИИ. Только так можно учесть множество таких факторов, как оперативная обстановка, способность системы работать автономно во враждебной среде с учетом воздействия, оказываемого на систему, параметры системы и взаимодействие человека и машины.

*Научно-исследовательские и опытно-конструкторские работы имеют ключевое значение для ответственного внедрения приложений искусственного интеллекта в военной области*

Нидерланды убеждены в важности научно-исследовательских и опытно-конструкторских работ. Государства должны адекватно оценивать, действуют ли их приложения ИИ так, как они были задуманы, и могут ли они быть развернуты в той или иной конкретной ситуации. Это особенно необходимо во время боевых действий и в других ситуациях высокого риска. Научно-исследовательские и опытно-конструкторские работы в целом, а также проверенные и надежные процедуры тестирования, оценки, проверки и проведения приемочных испытаний для конкретных приложений ИИ позволяют выявлять потенциальные проблемы и устранять или смягчать их до внедрения таких приложений. Кроме того, важно обеспечить надлежащую подготовку личного состава и изучение приложений ИИ до начала их применения, чтобы понимать возможности и ограничения таких приложений. Это особенно важно с учетом стремительного развития технологий, связанных с применением ИИ, и того факта, что его применение становится все менее затратным.

*Международное регулирование военного применения искусственного интеллекта должно быть гибким, инклюзивным и реалистичным*

Что касается международного регулирования в сфере применения ИИ в военной области, то Нидерланды признают необходимость принятия гибкого, сбалансированного и реалистичного подхода. Во-первых, система регулирования должна быть гибкой, чтобы успевать за стремительным развитием технологий и характера боевых действий. Во-вторых, сторонам необходимо работать над общим пониманием концепции применения ИИ в военной области, а также возможностей, рисков и потенциальных решений, сопутствующих такому применению. Это потребует всеохватного глобального диалога и активного участия всех заинтересованных сторон, включая государства, научные учреждения, гражданское общество и отраслевые структуры. В-третьих, государствам следует сделать упор на установлении гарантий ответственного применения ИИ в военной области, например путем уделения особого внимания таким вопросам,

как обеспечение отслеживаемости или объяснимости. В-четвертых, международное регулирование в сфере внедрения военного ИИ должно учитывать различные взгляды государств на вопросы регулирования. В рамках существующих правовых обязательств международное регулирование применения ИИ в военной области не должно ограничивать возможности государств по внедрению новаторских разработок.

### **Обсуждение вопроса об автономных системах вооружений**

Поскольку ИИ обладает значительным потенциалом в плане управления автономными системами вооружений, просматриваются четкие параллели между более широким обсуждением вопроса о его применении в военной области и обсуждением вопроса о регулировании автономных систем вооружений. По мнению Нидерландов, международные обсуждения по обеим темам взаимно дополняют и подкрепляют друг друга.

### **Новая Зеландия**

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Настоящий материал, представленный Новой Зеландией в национальном качестве, является ответом на вербальную ноту Управления по вопросам разоружения от 12 февраля 2025 года и должен рассматриваться в сочетании с ответом Новой Зеландии на вербальную ноту Управления от 1 февраля 2024 года<sup>1</sup>.

### **Позиция Новой Зеландии по вопросу о применении искусственного интеллекта в военной области**

Новая Зеландия признает, что потенциальные и уже существующие виды применения искусственного интеллекта (ИИ) в военной области будут иметь далеко идущие и многоплановые последствия.

Пока неясно, какими будут характер и масштабы многих из этих воздействий, но ИИ уже применяется некоторыми военными структурами для решения военных задач широкого спектра, в том числе для разведки, планирования, тылового обеспечения, навигации и связи. Несмотря на наличие некоторых рисков, применение ИИ в военной области может дать пользователям значительные преимущества, в том числе в плане скорости, эффективности, точности и ситуационной осведомленности. Как и вооруженные силы других стран, Силы обороны Новой Зеландии намерены использовать потенциал ИИ для повышения своего боевого потенциала и поддержания оперативной совместимости со своими партнерами.

Мы подтверждаем пункт 1 резолюции 79/239 Генеральной Ассамблеи, а именно, положение о том, что «нормы международного права, включая Устав Организации Объединенных Наций, международное гуманитарное право и международное право прав человека, применяются к регулируемым ими вопросам, которые возникают на всех этапах жизненного цикла средств искусственного интеллекта, в том числе к системам, функционирование которых в военной области обеспечивает искусственный интеллект». Помимо требующих исполнения правовых обязательств, на протяжении всего жизненного цикла средств ИИ,

<sup>1</sup> URL: [www.mfat.govt.nz/assets/Peace-Rights-and-Security/Disarmament/New-Zealand-submission-to-the-UN-Secretary-General-on-autonomous-weapon-systems.pdf](http://www.mfat.govt.nz/assets/Peace-Rights-and-Security/Disarmament/New-Zealand-submission-to-the-UN-Secretary-General-on-autonomous-weapon-systems.pdf).

применяемых в военной области, следует учитывать и соответствующие этические нормы.

Новая Зеландия признает, что ИИ имеет отношение к разработке и применению некоторых систем вооружений, например, для повышения степени автономности. Позиция Новой Зеландии в отношении автономных систем вооружений подробно изложена в ее ответе на вербальную ноту Управления по вопросам разоружения от 1 февраля 2024 года.

Не исключено, что ИИ может использоваться для разработки оружия массового уничтожения. Биологическое и химическое оружие однозначно запрещено нормами международным правом, и Новая Зеландия подтверждает, что критерий общей цели, сформулированный как в Конвенции о запрещении разработки, производства и накопления запасов бактериологического (биологического) и токсинного оружия и об их уничтожении, так и в Конвенции о запрещении разработки, производства, накопления и применения химического оружия и о его уничтожении, будет применяться и к использованию ИИ для разработки такого оружия, что означает, в частности, что ИИ не должен использоваться в таких целях. Кроме того, как отмечают государства — участники Договора о запрещении ядерного оружия, включая Новую Зеландию, в период до ликвидации ядерного оружия и систем его доставки и построения мира, свободного от ядерного оружия, существенно важное значение имеет сохранение значимого человеческого контроля над ними.

### **Существующие и новые нормативные предложения**

Достижение общего понимания и выработка норм являются важными аспектами содействия ответственному военному применению ИИ. В 2024 году Новая Зеландия, как и многие другие страны, присоединилась к Политической декларации об ответственном применении ИИ и автономных систем в военных целях, инициированной Соединенными Штатами. В этой декларации указывается, что «военное применение ИИ может и должно быть этичным и ответственным и должно способствовать укреплению международной безопасности». Кроме того, Новая Зеландия приняла участие в саммите по ответственному использованию искусственного интеллекта в военной сфере.

Новая Зеландия считает полезными многосторонние обсуждения, в том числе в рамках Организации Объединенных Наций, посвященные разработке и согласованию норм, касающихся применения ИИ в военной области. Участие в этих обсуждениях представителей негосударственных заинтересованных сторон, в том числе гражданского общества, международных и региональных организаций, а также отраслевых структур, имеет важное значение на протяжении всего процесса.

### **Норвегия**

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Норвегия приветствует эту возможность представить свои мнения о возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений в соответствии с резолюцией 79/239 Генеральной Ассамблеи, озаглавленной «Применение искусственного интеллекта в военной области и его последствия для международного мира и безопасности».

Как отмечается в концептуальной записке Генерального секретаря от июля 2023 года «Новая повестка дня для мира», ИИ — это одновременно и ценная, и разрушающая устоявшийся уклад технология, которая все больше интегрируется в широкий спектр систем гражданского, военного и двойного назначения. Все более широкое применение ИИ в сочетании с его быстрой масштабируемостью, отсутствием транспарентности и темпами инноваций представляет собой потенциальные риски для международного мира и безопасности и проблемы в области регулирования.

В качестве последовательного поборника международного права, многосторонности и ответственных инноваций в оборонном секторе Норвегия поддерживает усилия по содействию выработке общего понимания, усилению механизмов управления и выработке адекватных правил, регулирующих применение ИИ в военной области. Как минимум необходимо обеспечить ответственный подход к разработке, внедрению и применению приложений ИИ в военной области на протяжении всего их жизненного цикла и их соответствие применимым нормам международного права, в частности международного гуманитарного права.

Важно отметить, что в резолюции [79/239](#) Генеральная Ассамблея подтвердила применимость норм международного права, включая Устав Организации Объединенных Наций, международное гуманитарное право и право прав человека, в процессе применения ИИ в военной области и подчеркнула важность ответственного человекоцентричного применения ИИ.

ИИ в качестве обеспечивающей технологии обладает огромным потенциалом для преобразования всех аспектов военного дела, включая закупки, аппаратно-программные средства, ведение операций, командование и управление, стратегическую связь, наблюдение, разведку, обучение личного состава, управление информацией и тыловое обеспечение. Применение ИИ в военной области создает предсказуемые и непредсказуемые возможности и риски как на тактическом, так и на стратегическом уровне. В качестве технологии общего назначения ИИ представляет собой фактор повышения боевой эффективности, способный изменить способы ведения войны. Дополнительные трудности создает взаимопроникновение технологий искусственного интеллекта, нейротехнологии, синтетической биологии и квантовых вычислений.

Основополагающее значение имеет ответственный подход к разработке, внедрению, применению и регулированию ИИ в соответствии с базовыми этическими принципами, при строгом соблюдении обязательств государств по международному праву, включая международное гуманитарное право и право прав человека, с уделением первостепенного внимания выявлению и уменьшению рисков.

В Стратегии развития искусственного интеллекта в оборонном секторе Норвегии (2023 год) излагаются перечисленные ниже ключевые области, в которых ИИ может внести конструктивный вклад в развитие других направлений помимо смертоносных автономных систем вооружений.

- **Повышение ситуационной осведомленности и поддержка принятия решений.** Применение ИИ является не только обеспечивает различные возможности, но и является необходимым фактором в сфере разведки, наблюдения и рекогносцировки, поскольку большие и все возрастающие объемы данных невозможно проанализировать вручную. ИИ может использоваться для выявления релевантных данных, например путем их предварительной обработки, автоматического перевода или распознавания искомых объектов на изображениях, выявления аномалий и повторов, а также

перекрестной проверки информации в целях обнаружения попыток дезинформации. Прогресс в этой области может способствовать повышению эффективности и точности операций и уменьшению числа жертв.

- **Киберзащита.** Внедрение цифровых технологий и все более широкая опора на информационно-коммуникационные технологии не только создают преимущества, но и приносят факторы уязвимости. Цифровое пространство предоставляет злоумышленникам возможность осуществлять несанкционированные доступы к данным, вести шпионскую и подрывную деятельность, а также проводить кампании по оказанию влияния. ИИ способен поддерживать потенциал оборонного сектора по выявлению и отслеживанию цифровых угроз, сообщению о них, противодействию им и их пресечению. Применение ИИ позволяет, в частности, ускорять процесс получения более полного представления о целях и сложных взаимосвязях, осуществлять сбор информации из соответствующих источников и оптимизировать использование аналитических данных. Знания и опыт касательно способов применения ИИ для создания цифровых угроз имеют существенно важное значение для того, чтобы иметь возможность выявлять и предотвращать цифровые атаки в будущем. Поэтому ИИ должен стать одним из центральных элементов в сфере дальнейшего развития защиты сектора от цифровых угроз с помощью существующих и будущих инструментов.
- **Тыловое обеспечение.** Успех и эффективность военных операций зависят от эффективного тылового обеспечения. Оптимизация тылового обеспечения благодаря применению систем с использованием ИИ позволяет расширить боевые возможности и повысить боеготовность. ИИ уже широко применяется в гражданском секторе логистики. Многие из таких гражданских видов применения можно легко адаптировать для военного сектора.
- **Вспомогательная деятельность.** Многие виды обеспечения военных действий можно усовершенствовать и оптимизировать с помощью ИИ. К их числу относятся задачи по поддержке и наращиванию боевых возможностей, например эксплуатация и техническое обслуживание материальной части, приобретение материальной части и сооружений, управление ими и их утилизация, комплектование и подготовка личного состава и управление делами личного состава, а также предоставление таких общих услуг, как бухгалтерский учет и ведение архивов. ИИ способен упрочить вспомогательную деятельность путем более эффективного использования данных для анализа и поддержки принятия решений, автоматизации задач и расширения возможностей обработки информации и знаний. Это позволило бы перейти к модели прогнозируемого технического обслуживания, улучшить информационные потоки, внедрить новые и более совершенные системы поддержки управления делами личного состава и усовершенствовать моделирование тенденций в сфере расходов на материальную часть и сооружения. Поэтому успешное внедрение технологии ИИ в рамках вспомогательной деятельности могло бы привести к сокращению затрат времени и повышению эффективности.

Кроме того, применение ИИ в военной области способно повысить уровень соблюдения норм международного гуманитарного права и поддержать усилия по защите гражданского населения и гражданских объектов в вооруженных конфликтах. ИИ может быть полезен для миростроительства и мероприятий по поддержанию мира, а также может способствовать наращиванию потенциала в сфере проверки и мониторинга в рамках режимов контроля над вооружениями, разоружения и других режимов соблюдения.



Применение ИИ в военной области также порождает беспрецедентные проблемы. ИИ присущи факторы уязвимости, которые могут иметь непреднамеренные последствия и могут ослабить полноценный контроль, ответственность и подотчетность оператора. Использование глубокого обучения может усложнить задачи понимания, объяснения и предсказания поведения моделей ИИ. Например, недостаточный уровень понимания может сделать динамику эскалации конфликта менее прозрачной и предсказуемой.

Необходимо обеспечить эффективные гарантии сохранения человеком значимого контроля и надзора за разработкой, внедрением и применением ИИ. Это приобретает особенно важное значение в тех случаях, когда приложения ИИ используются для выполнения функций, приближающихся к ведению боевых операций и применению силы, например, системы поддержки принятия решений. Ответственность за применение и последствия применения военного ИИ всегда должна лежать на человеке.

Системы ИИ могут быть очень чувствительны к качеству и репрезентативности данных, используемых в процессе обучения. Возможные предвзятость, ложный или неполный характер обучающих данных могут привести к созданию моделей, генерирующих неточные или дискриминационные результаты. Предвзятость автоматизации может вызывать чрезмерное доверие пользователя к выходным данным системы.

Использование в киберпространстве средств реагирования с высокой степенью автоматизации или автономных средств реагирования, особенно средств, не предусматривающих надлежащего участия оператора в контуре управления, может вызвать непреднамеренные ответные действия и быструю эскалацию.

Все более широкая опора на кибертехнологии при решении задач, которые раньше выполнялись вручную или с помощью базовых средств автоматизации, также сопряжена с риском злонамеренного использования факторов уязвимости этих технологий. Все более широкое использование коммерческих систем вызывает опасения по поводу зависимости от внешних поставщиков, утраты контроля над обновлениями и других факторов уязвимости, связанных с фирменными системами.

Выше были перечислены лишь некоторые примеры потенциальных рисков, связанных с применением ИИ в военной области. Существует и множество неизвестных факторов. В военном контексте эти факторы в совокупности или сами по себе могут сорвать достижение целей боевого задания и создать фундаментальные правовые, этические, гуманитарные и военные риски.

В Стратегии развития искусственного интеллекта в оборонном секторе Норвегии (2023 год) также изложены перечисленные ниже ключевые принципы ответственного развития и применения ИИ.

- **Законность.** Приложения ИИ надлежит разрабатывать и применять в соответствии с нормами международного права, включая международное гуманитарное право и право прав человека. При изучении, разработке, приобретении или принятии на вооружение новых видов оружия, средств или методов ведения войны, использующих ИИ, каждое государство должно определить, подпадает ли их применение, при некоторых или при всех обстоятельствах, под запрещения, содержащиеся в нормах международного гуманитарного права или в каких-либо других нормах международного права, применяемых к этому государству.

- **Ответственность и подотчетность.** Необходимо обеспечить ответственность и подотчетность человека в связи с применением ИИ. Необходимо четко определить полномочия по принятию решений на применение системы ИИ и ответственность за ее фактическое применение.
- **Объяснимость, понятность, отслеживаемость.** Приложения ИИ должны быть в достаточной степени объяснимыми, понятными, транспарентными и отслеживаемыми.
- **Подготовка личного состава.** Операторы ИИ должны иметь необходимую подготовку, чтобы понимать поведение приложений ИИ, включая способы выявления аномального поведения.
- **Надежность, безопасность и защищенность.** Приложения ИИ должны иметь четкие и точно сформулированные сферы применения. На протяжении всего жизненного цикла приложений ИИ в рамках соответствующих сфер применения необходимо тестировать и проверять их отказоустойчивость, надежность и безопасность. Приложения ИИ должны обладать достаточным уровнем безопасности и иметь защиту от цифровых угроз.
- **Контроль.** Необходимо обеспечить значимый человеческий контроль. Системы ИИ должны предусматривать соответствующий назначению таких систем интерфейс для взаимодействия оператора и машины, обеспечивающий возможность выявления и смягчения непреднамеренных последствий, а также средства для ввода коррекций, если система выходит на штатный режим.

Международному сообществу необходимо углубить диалог по вопросу о военных приложениях ИИ и последствиях их применения для мира и безопасности, в том числе о мерах по обеспечению ответственного применения ИИ в военной области. Особое внимание следует уделять системам поддержки боевых операций, включая применение ИИ для обеспечения ситуационной осведомленности и поддержки принятия решений, где нежелательные результаты и поведение приложений ИИ, а также утрата значимого человеческого контроля могут иметь особенно пагубные последствия. Необходимо также рассмотреть вопросы применения ИИ в гибридных войнах, включая, в частности, его применение в кибероперациях, радиоэлектронной борьбе и информационных операциях.

Норвегия решительно выступает за укрепление международного сотрудничества в сфере обмена информацией и наращивания потенциала. Путем создания общей базы знаний государства будут способствовать общему пониманию, устранению пробелов, повышению степени транспарентности и укреплению доверия. С этой целью Норвегия намеревается поощрять разработку и публикацию национальных стратегий и директивных документов, касающихся военного применения ИИ. Необходимо уделять внимание уменьшению рисков и принятию мер укрепления доверия.

Своевременное создание надлежащей международной системы регулирования ИИ, которая позволяла бы гибко реагировать на стремительный технологический прогресс, может способствовать предотвращению гонки вооружений, вызванной развитием технологий, и обеспечить поддержку инноваций в сфере обеспечения глобальной безопасности.

## Пакистан

[Подлинный текст на английском языке]  
[9 апреля 2025 года]

Стремительное развитие и интеграция технологий искусственного интеллекта (ИИ) в военной области способны коренным образом изменить характер военных действий. ИИ все шире интегрируется в военные операции благодаря применению в автономных системах вооружений, командованию и управлению, системах поддержки принятия решений, разведке, наблюдению и рекогносцировке, подготовке личного состава, тыловом обеспечении, и ведении кибер-/информационных войн. Такое развитие повышает боеспособность, но при этом порождает значительные риски для международного мира и безопасности.

### **Проблемы, связанные с применением искусственного интеллекта в военной области**

#### *Стратегические риски: взаимодействие с ядерным оружием*

Задействование ИИ в системах ядерного оружия создает стратегические риски, особенно в области ядерного командования, управления и связи. Когда средства ИИ включаются в политику строительства ядерных сил и стратегии их применения, это может вызвать просчеты, случайности и катастрофические последствия.

Концепция ядерного сдерживания в значительной мере зависит от рациональности, восприятия и принятия политических решений со стороны человека. Задействование ИИ потенциально устраняет или значительно ограничивает роль этих критически важных факторов участия человека, повышая риск эскалации, вызванной действиями автоматики или случайностью. Признавая эти проблемы, вызывающие глубокую озабоченность, некоторые государства публично обязались сохранять значимый контроль человека над принятием решений о применении ядерного оружия — принцип, который Пакистан поддерживает и настоятельно призывает принять все государства, обладающие ядерным оружием.

В регионах, где есть ядерное оружие, опора на системы поддержки принятия решений на основе ИИ и на полностью автономные системы обычных вооружений также может вызвать риски эскалации. Полное устранение контроля со стороны человека во время кризисов может затруднить контроль над масштабами и продолжительностью конфликтов. Автоматизация реагирования в нестабильных сценариях с высокими ставками, особенно в регионах с напряженной ядерной динамикой, может усугубить взаимодействие обычного и ядерного оружейных потенциалов и негативно сказаться на стратегической стабильности.

Использование ИИ для оценки данных и ведения разведки, наблюдения и рекогносцировки может породить ложное чувство уверенности у государств, рассматривающих возможность нанесения превентивных, дестабилизирующих контрсилловых ударов или ударов для подавления средств нанесения второго удара, что создает серьезные риски для региональной и глобальной стабильности.

#### *Операционные риски: ослабление участия человека*

Автономность военных операций, основанная на ИИ, может привести к понижению уровня контроля со стороны человека, что усложнит управление кризисными ситуациями. Когда военные действия разгоняются до «машинных скоростей», процесс принятия решений человеком сильно компрессируется, что

сокращает возможности для смягчения кризисных ситуаций и дипломатического вмешательства.

Люди могут чрезмерно доверять рекомендациям систем поддержки принятия решений, генерируемым ИИ, даже если они ошибочны или неполны, что приводит к принятию необъективных решений в результате использования автоматизации. Важнейшие военные решения могут чрезмерно зависеть от выводимых данных автоматизации, в результате чего командиры не обращают внимания на человеческую интуицию, не учитывают контекст или не проявляют осторожность, что может привести к непреднамеренной эскалации конфликтов.

Возможности, основанные на использовании ИИ, которые стимулируются повышением операционной эффективности и борьбой за решающее преимущество, могут привести к увеличению вероятности их применения, что снизит порог развязывания вооруженного конфликта. В условиях кризиса низкий порог применения силы будет иметь крайне дестабилизирующий эффект.

#### *Технические риски*

Применение ИИ в военных целях может породить факторы технической уязвимости, в том числе алгоритмическую предвзятость, искажение данных и подверженность кибератакам. Конфликты могут возникать из-за сбоев в работе систем раннего оповещения или манипуляций с ними, а также из-за атак, направленных на искажение данных. Средства ИИ часто работают по принципу «черного ящика», в результате чего решения не являются прозрачными или объяснимыми, что затрудняет валидацию и отчетность. Такие факторы уязвимости могут привести к непредсказуемым результатам, системным сбоям и значительным рискам для операционной надежности. Возможности ИИ, протестированные в одной среде с конкретными наборами данных, могут ненадежно работать в совершенно других средах с более сложной динамикой.

#### *Нормативные, правовые и этические риски*

Применение ИИ в военной области создает этические, нормативные и правовые проблемы, особенно в том, что касается соблюдения норм международного гуманитарного права. Суть международного гуманитарного права основывается главным образом на человеческом суждении, свободе выбора и принятии решений с учетом контекста, то есть на качествах, которые трудно воспроизвести системам ИИ по определению. Делегирование автономным системам таких критически важных функций, как выбор и поражение цели, включая принятие решений о применении смертоносной силы, сопряжено с риском нарушения основных принципов международного гуманитарного права, касающихся обеспечения избирательности и соразмерности, принятия мер предосторожности при нападении и учета военной необходимости. Системы ИИ, которые приводят к непредсказуемым, ненадежным или необъяснимым результатам, еще больше усложняют задачу соблюдения норм международного гуманитарного права, потенциально приводя к незаконному или непреднамеренному ущербу.

Кроме того, отсутствие прямого участия человека в принятии решений или чрезмерная зависимость от систем поддержки принятия решений, основанных на ИИ, поднимает важнейшие вопросы отчетности и ответственности, что делает установление виновности и ответственности в случаях незаконных или противоправных действий чрезвычайно сложной задачей. Если что-то пойдет не так, то командиры могут переложить ответственность на ИИ, что усложнит юридическую ответственность и расследование потенциальных военных преступлений.

Этические проблемы возникают также из-за делегирования автономным системам полномочий принимать решения о жизни и смерти, что потенциально снижает уровень сострадания, нравственной аргументации и здравого смысла, тем самым повышая риск неоправданного насилия и жертв среди гражданского населения.

#### *Риски распространения и риски для глобальной безопасности*

Распространение военных технологий ИИ создает значительные риски для международной безопасности. Распространение передовых возможностей ИИ, в частности автономного оружия, может спровоцировать новую гонку вооружений и дестабилизировать обстановку в области региональной и глобальной безопасности. Легкость распространения и возможность приобретения негосударственными субъектами еще больше усиливают эти опасения.

#### **Предлагаемый международный ответ: центральная роль механизма Организации Объединенных Наций**

Технологии ИИ являются многоцелевыми, а их мирное использование является неотъемлемой частью достижения целей в области устойчивого развития. В то же время последствия применения ИИ в военной области носят сквозной характер и могут существенно повлиять на международный мир и безопасность, что потребует скоординированных международных ответных мер.

Пакистан признает ценность инициатив по регулированию ИИ за рамками Организации Объединенных Наций, но по-прежнему осознает их ограниченность, особенно в том, что касается всеобщего участия и официальной многосторонней легитимности. Эти инициативы могут дополнять усилия Организации Объединенных Наций, способствуя диалогу и мобилизации политической воли, однако их изолированная реализация чревата фрагментацией усилий. Поэтому дискуссии по вопросу о военном применении ИИ следует проводить на форумах Организации Объединенных Наций, с тем чтобы обеспечивать инклюзивность, легитимность и согласованную глобальную структуру, отражающую интересы всех государств.

По этим причинам Организация Объединенных Наций должна оставаться в центре любого международного реагирования. Разоруженческий механизм Организации Объединенных Наций должен играть центральную роль в разработке международной системы регулирования военного ИИ и предотвращения фрагментации нормативного ландшафта. Масштаб и новизна военных последствий применения ИИ требуют многогранного, целостного многостороннего реагирования. Универсальное членство в Организации Объединенных Наций однозначно делает ее идеальным форумом, в котором все государства — как развитые, так и развивающиеся — имеют право голоса.

Одного форума или инструмента будет недостаточно. Требуется структурированная стратегия, использующая многочисленные разоруженческие органы Организации Объединенных Наций, в рамках которых каждый форум рассматривает проблему со своего уникального ракурса и на базе своего мандата на взаимодополняющей основе. Мы предлагаем задействовать все соответствующие форумы — от Генеральной Ассамблеи и ее Первого комитета до Комиссии по разоружению, Конференции по разоружению и Конвенции о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие. Такой подход позволил бы всесторонне учитывать стратегические, гуманитарные, правовые и технические аспекты, избегая пробелов и избыточности. Работа каждого форума должна учитываться в работе других,

создавая синергию для достижения общей цели, которая заключается в уменьшении рисков, связанных с военным применением ИИ, при сохранении мирного использования ИИ.

#### *Конференция по разоружению*

Конференции по разоружению следует уделять первостепенное внимание рассмотрению стратегических рисков, связанных с военным ИИ, особенно в ядерной области, что непосредственно согласуется с пунктами 1 и 2 ее повестки дня («Прекращение гонки ядерных вооружений и ядерное разоружение» и «Предотвращение ядерной войны, включая все связанные с этим вопросы»). В 2023 году Пакистан предложил включить в повестку дня Конференции по разоружению новый пункт повестки дня по этому вопросу ([CD/2397](#)).

В рамках этого нового пункта повестки дня Конференции по разоружению следует учредить вспомогательный орган или специальную группу с конкретным мандатом по изучению рисков военного применения ИИ для стабильности, составления оценки того, каким образом оно усиливает ядерные риски, и проведения переговоров о конкретных мерах. Эти меры могли бы включать:

- принятие обязательства по сохранению человеческого контроля и недопущению подмены человеческого суждения в решениях, касающихся применения ядерного оружия;
- запрещение использования средств ИИ для манипулирования данными или нацеливания на системы ядерного командования, управления и связи;
- разработка мер по сдерживанию развертывания и применения определенных средств ИИ, которые могут инициировать превентивные удары и порождать ядерные риски, вызывающие эскалацию.

Конференция по разоружению уникально подходит для проведения таких дискуссий, объединяя все значимые в военном отношении государства на равной основе и действуя на основе консенсуса, тем самым защищая жизненно важные интересы безопасности всех государств. Рассмотрение этой проблемы могло бы активизировать работу Конференции по разоружению, продемонстрировав ее способность реагировать на новые и нарождающиеся угрозы.

#### *Комиссия по разоружению*

Обладая универсальным членским составом и совещательным мандатом, Комиссия по разоружению идеально подходит для разработки практических руководящих принципов и рекомендаций по ответственному применению ИИ в военных целях. В прошлом Комиссия по разоружению разработала аналогичные руководящие принципы (например, меры укрепления доверия в 1988 году и региональные подходы к разоружению в 1993 году).

В рамках своей рабочей группы II Комиссия по разоружению могла бы разработать руководящие принципы и рекомендации по мерам укрепления доверия и безопасности, связанным с применением ИИ в военных целях, как на глобальном, так и на региональном уровнях. Ключевыми элементами могли бы стать подтверждение нормативных основ, составление рекомендаций по оперативным и техническим мерам для смягчения рисков, разработка стратегий уменьшения рисков, связанных с применением ИИ в военных целях, и решение проблем, связанных с распространением, при обеспечении равного доступа к использованию ИИ в мирных целях.

*Первый комитет Генеральной Ассамблеи Организации Объединенных Наций*

Первому комитету Генеральной Ассамблеи следует институционализировать регулярные оценочные доклады Генерального секретаря Организации Объединенных Наций и вести каталог технологических разработок в области военного применения ИИ и связанных с ними рисков на основе добровольной информации, которой делятся государства-члены. Эти периодические оценки позволили бы получать достоверное представление о развивающихся возможностях, предоставляя своевременную информацию и способствуя принятию обоснованных мер реагирования в сфере международной политики.

При рассмотрении таких докладов Первый комитет мог бы провести специальные прения по вопросам ИИ и, возможно, учредить рабочую группу открытого состава в рамках Генеральной Ассамблеи, если это потребуется для обсуждения более институциональной платформы, например регистра Организации Объединенных Наций по применению ИИ в военных целях (хотя на данный момент представляется предпочтительным использовать существующие форумы).

Эти доклады могли бы также определять области, в которых формируется консенсус или требуется дальнейшая работа, определяя повестки дня таких форумов, как Конференция по разоружению, Комиссия по разоружению и Конвенция по конкретным видам обычного оружия.

*Конвенции по конкретным видам обычного оружия*

Группа правительственных экспертов, созданная в рамках Конвенции по конкретным видам обычного оружия, по-прежнему играет существенно важную роль в рассмотрении гуманитарных, этических и правовых последствий применения смертоносных автономных систем вооружений. Одним из ее достоинств является ее инклюзивный характер (привлечение гражданского общества и Международного комитета Красного Креста в качестве наблюдателей).

Пакистан высоко оценивает работу, проделанную Группой правительственных экспертов с 2017 года, в частности 11 руководящих принципов, принятых в 2019 году. Однако прогресс в рамках Конвенции протекает медленными темпами и касается в основном принципов, а не конкретных положений. Пакистан согласен с оценками, согласно которым в дискуссиях в рамках Конвенции уделяется «недостаточное и все меньшее внимание» аспектам безопасности оружия, в котором используется ИИ, что подчеркивает необходимость принятия дополнительных мер на Конференции по разоружению и других форумах. Тем не менее, на гуманитарном направлении Группе правительственных экспертов, созданной в соответствии с Конвенцией, следует продолжать и активизировать свою работу.

Пакистан выступает за завершение переговоров о юридически обязывающем протоколе к Конвенции, запрещающем смертоносные автономные системы вооружений, действующие без контроля человека или неспособные соответствовать нормам международного гуманитарного права. Нынешний мандат Группы правительственных экспертов позволяет государствам-членам разрабатывать элементы такого документа для представления на седьмой обзорной конференции высоких договаривающихся сторон Конвенции и впоследствии потенциально начать официальные переговоры.

## Заключение

Пакистан подчеркивает необходимость принятия скоординированных, инклюзивных международных мер для смягчения существенных военных рисков, связанных с ИИ. Он выступает за принятие такого подхода к регулированию, который поддерживает баланс интересов безопасности и развития, обеспечивает стабильность и в то же время способствует благотворному развитию ИИ. Благодаря структурированной стратегии с участием нескольких форумов в рамках Организации Объединенных Наций международное сообщество может установить надежные нормативные барьеры, поддержать международную безопасность и сохранить справедливый, недискриминационный доступ к использованию ИИ в мирных целях.

## Перу

[Подлинный текст на испанском языке]  
[11 апреля 2025 года]

В пункте 7 резолюции [79/239](#) Генеральной Ассамблеи Организации Объединенных Наций от 24 декабря 2024 года, за которую голосовала Перу, к Генеральному секретарю обращена просьба:

запросить мнения государств-членов и государств-наблюдателей о возможностях и проблемах в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта в военной области, с уделением особого внимания областям помимо смертоносных автономных систем вооружений, и представить Генеральной Ассамблее на ее восьмидесятой сессии для дальнейшего обсуждения государствами субстантивный доклад, содержащий обобщение этих мнений и каталог существующих и новых нормативных предложений, а также приложение с изложением этих мнений.

В этой связи Перу представляет ниже некоторые элементы своей позиции, чтобы внести вклад в подготовку вышеупомянутого доклада Генерального секретаря.

### I. Актуальность применения искусственного интеллекта в военной области

Перу отмечает бурное и динамичное развитие новых технологий, предназначенных для применения в военной области, в частности потенциальные виды применения искусственного интеллекта (ИИ). Мы внимательно отслеживаем развитие событий в этой области, в том числе потенциальные виды применения ИИ, способные изменить военные операции — от применения автономных безоператорных систем до использования систем поддержки принятия решений, — и считаем необходимым развивать постоянный многосторонний диалог в целях разработки принципов, обеспечивающих этическое и ответственное использование таких технологий.

Поскольку ИИ может быть интегрирован как в системы вооружения, так и в системы поддержки военных операций, Перу считает необходимым рассмотреть проблемы и опасения, связанные с его применением, с гуманитарной, правовой, технологической и этической точек зрения, включая риски, связанные с алгоритмической предвзятостью. Такие опасения усиливаются последствиями, которые может иметь применение этой технологии для международной стабильности и безопасности.



Это тем более тревожно с учетом риска последствий применения ИИ в связи с ядерным оружием и другими видами оружия массового уничтожения. В этой связи крайне важно укрепить принцип значимого человеческого контроля.

## **II. Мнения**

### *Соблюдение норм международного права*

Разработка, внедрение и применение технологий с поддержкой ИИ в военной области должны полностью соответствовать нормам международного права, включая международное право прав человека и международное гуманитарное право, а также основополагающим принципам, закрепленным в Уставе Организации Объединенных Наций.

Поэтому любые разработки в сфере регулирования ИИ в военной области должны обеспечивать его ответственное и этичное применение, что также гарантирует нераспространение военных технологий с поддержкой ИИ и равноправный доступ к знаниям и технологическим возможностям.

Это необходимо для того, чтобы в процессе применения ИИ во всех случаях обеспечивались соблюдение человеческого достоинства, защита гражданского населения и международная стабильность и мир.

### *Признание преимуществ и рисков*

ИИ открывает ценные возможности для более глубокого понимания оперативной обстановки и, следовательно, для обеспечения соблюдения норм международного гуманитарного права и защиты гражданского населения и гражданских объектов.

Вместе с тем применение ИИ может породить предсказуемые и непредсказуемые риски в военной области, например вследствие алгоритмической предвзятости, конструктивных недостатков, неправомерного или злонамеренного применения и т.д. Кроме того, ИИ может влиять на сложную региональную и глобальную динамику, усиливая риски эскалации, просчетов, понижения порога конфликта или развязывания гонки вооружений.

### *Ответственное развитие*

Применение ИИ в военной области должно способствовать укреплению мира и защите гражданского населения, а технологические достижения должны дополнять, а не подменять собой человеческие возможности.

В соответствии с принципами, применимыми к автономным системам вооружений, в процессе применения ИИ в военной области необходимо гарантировать, что ответственность и подотчетность никогда не будут перекладываться на машины. В этой связи нужно особо отметить необходимость сохранения значимого человеческого контроля при принятии любых решений, связанных с применением силы.

Необходимо обеспечить комплексный подход ко всем рискам и проблемам, связанным со средствами ИИ, на протяжении всего их жизненного цикла.

Обеспечение механизмов контроля и гарантий во избежание неправомерного применения этой технологии в военной области не должно препятствовать исследованиям, разработкам, экспериментам и инновациям, связанным с ИИ, в других сферах.

### *Внедрение и транспарентность*

Задачей первостепенной важности является определение стратегий, принципов, стандартов и норм, а также выработка национальной политики и правовых основ для обеспечения ответственного применения ИИ в военной области.

Кроме того, необходимо особо отметить важность разработки мер укрепления доверия и уменьшения рисков, а также создания механизмов обмена передовым опытом в интересах обеспечения транспарентности и сотрудничества между государствами.

### *Формат обсуждений*

Существенно важное значение имеет продолжение диалога на глобальном, региональном и межгосударственном уровнях для обсуждения вопросов разработки мер, обеспечивающих ответственное применение ИИ в военной области.

При этом необходимо обеспечивать всестороннее участие в обсуждении этой темы, в том числе позиций государств, в частности развивающихся стран, а также вклада других заинтересованных сторон, таких как отраслевые и научные круги, гражданское общество, региональные и международные организации.

Важно учитывать, что разные государства и регионы находятся на разных этапах деятельности по внедрению средств ИИ в военной области и находятся в разных ситуациях в плане безопасности.

В этой связи подчеркивается важность содействия наращиванию потенциала в развивающихся странах и укрепления международного сотрудничества в целях сокращения существующего отставания и расширения их участия в обсуждении вопросов использования этой технологии.

### **Участие Перу в международных дискуссиях**

#### *Саммиты по ответственному использованию искусственного интеллекта в военной сфере*

- Перу участвовала в саммитах 2023 и 2024 годов и в соответствующем региональном практикуме
- Перу одобрила заключительную декларацию Саммита по ответственному использованию искусственного интеллекта в военной сфере 2024 года («План действий»)

Политическая декларация об ответственном применении искусственного интеллекта и автономных систем в военных целях:

- Перу в качестве наблюдателя участвовала в первом пленарном заседании по линии этой инициативы и впоследствии официально заявила о своем присоединении к ее участникам

Саммит действий по искусственному интеллекту — переговоры по военным вопросам (Париж, 2025 год):

- Перу на высоком уровне приняла участие в этом саммите и подписала Парижскую декларацию о сохранении человеческого контроля в оружейных системах с поддержкой ИИ.

## Республика Корея

[Подлинный текст на английском языке]

[11 апреля 2025 года]

В качестве обеспечивающей технологии искусственный интеллект (ИИ) способен кардинально изменить множество аспектов военного дела — от принятия решений и сбора разведывательных данных до тылового обеспечения, наблюдения и систем командования и управления. Бурное развитие ИИ сопровождается ростом интереса государств к применению этой технологии в военной области.

Все более широкое применение средств ИИ и систем с поддержкой ИИ в рамках военных операций открывает возможности и порождает проблемы, особенно для международного мира и безопасности. Это затрагивает важные вопросы с гуманитарной, правовой, технологической и этической точек зрения.

Для целей настоящего представления изложенные ниже мнения касаются исключительно вопросов помимо смертоносных автономных систем вооружений.

### **Возможности, связанные с применением искусственного интеллекта в военной области**

Средства ИИ и системы с поддержкой ИИ, в том числе используемые в системах разведки, наблюдения и рекогносцировки и поддержки принятия решений, позволяют повысить уровень ситуационной осведомленности, точности и безошибочности, а также эффективности благодаря обработке больших объемов данных, поддержки оптимизации и генерирования аналитических оценок. Такие средства и системы могут способствовать поддержанию и укреплению международного мира и безопасности.

1. *Повышение эффективности соблюдения норм международного гуманитарного права и содействие обеспечению защиты гражданского населения и гражданских объектов в вооруженных конфликтах*

Системы ИИ, используемые для разведки, наблюдения и рекогносцировки, а также поддержки принятия решений, могут повысить эффективность применения основополагающих принципов международного гуманитарного права (избирательность, соразмерность и предосторожность при нападении) путем обеспечения более точной оценки ситуации на театре военных действий и улучшения ситуационной осведомленности. Благодаря доступу к своевременной и достаточной информации ИИ может способствовать проведению различий между комбатантами и некомбатантами, а также оценке потенциального сопутствующего ущерба. Путем повышения осведомленности о ситуации на театре военных действий, в том числе о присутствии гражданского населения, ИИ помогает определить необходимость и целесообразность принятия мер предосторожности для защиты гражданского населения и гражданской инфраструктуры.

2. *Поддержка операций по поддержанию мира*

ИИ может поддерживать мониторинг соглашений о прекращении огня и мирных договоренностей. Кроме того, ИИ может способствовать созданию механизмов раннего предупреждения для выявления потенциальных нарушений, повышая эффективность и безопасность миссий. Республика Корея приступила к осуществлению экспериментального проекта «умного лагеря» в подразделении «Ханбит» Миссии Организации Объединенных Наций в Южном Судане, направленного на повышение безопасности и эффективности и расширение

оперативных возможностей миротворческих лагерей Организации Объединенных Наций путем применения ИИ и других новейших технологий.

3. *Наращивание потенциала проверки и мониторинга для режимов контроля над вооружениями и соблюдения*

ИИ может расширять возможности международных механизмов проверки для мониторинга соблюдения соглашений о контроле над вооружениями и нераспространении. Международное агентство по атомной энергии может использовать ИИ для повышения эффективности гарантийных процессов, в частности тех, которые связаны с классификацией данных, поиском закономерностей и выявлением выбросов в данных. Кроме того, системы с поддержкой ИИ могут способствовать выявлению ранних признаков применения химического или биологического оружия и все более изощренных способов обхода санкций, что способствовало бы укреплению международных режимов нераспространения.

Помимо описанных выше возможностей, ИИ может способствовать уменьшению таких стратегических рисков, как просчеты, недоразумения и непреднамеренная эскалация, путем совершенствования анализа поведения субъектов и наращивания потенциала упреждающего обнаружения и реагирования. Кроме того, средства ИИ могут способствовать наращиванию потенциала, в частности для укрепления киберзащиты, защиты важнейших национальных инфраструктурных объектов и борьбы с терроризмом.

**Проблемы, связанные с применением искусственного интеллекта в военной области**

Военное применение ИИ могло бы породить новые проблемы или усугубить уже существующие, если не обеспечить ответственный подход к его разработке, внедрению и применению.

Проблемы могут быть обусловлены техническими и эксплуатационными характеристиками ИИ. Например, заложенный в системах ИИ принцип «черного ящика» осложняет понимание того, как и почему генерируются конкретные выходные данные, что делает работу таких систем менее объяснимой и отслеживаемой. Конструктивные недоработки и непреднамеренная предвзятость в данных, алгоритмах или архитектуре системы могут вызывать сбои или приводить к выдаче результатов, отклоняющихся от намеченных целей. Чрезмерная опора на системы ИИ, например, смещение автоматизации или недостаточный уровень подготовки операторов могут вызывать проблемы, связанные с отсутствием надлежащего человеческого суждения и участия. Эти факторы могли бы повысить вероятность просчетов, недоразумений или непреднамеренной эскалации конфликта, что станет проблемой на пути обеспечения международного мира и безопасности.

Двойное назначение технологий ИИ могло бы усилить риск их неправомерного использования или злоупотребления ими со стороны безответственных субъектов со злым умыслом. Например, с помощью ИИ можно активизировать кампании дезинформации и кибератаки, в частности за счет подтасовки или подмены данных. Кроме того, безответственные субъекты могут использовать технологии ИИ для облегчения разработки нового химического или биологического оружия, что вызывает озабоченность по поводу распространения и усиливает риски для международного мира и безопасности.

### **Внедрение ответственного искусственного интеллекта в военной области**

Чтобы использовать преимущества и возможности ИИ, избегая при этом связанных с ним рисков и проблем, необходимо ответственно подходить к разработке, внедрению и применению в военной области средств ИИ и систем с его поддержкой на протяжении всего их жизненного цикла.

Республика Корея стремится обеспечивать и поощрять ответственное применение ИИ в военной области. Это предусматривает следующие ключевые принципы и меры:

- ИИ должен носить этический и человекоцентричный характер.
- Применение средств ИИ в военной области должно осуществляться с соблюдением применимых норм международного права, включая международное гуманитарное право и международное право прав человека.
- Ответственность и подотчетность в связи с применением ИИ в военной области и его последствиями должны лежать на человеке и никогда не могут перекладываться на машину.
- Безотказность и надежность приложений ИИ должны обеспечиваться путем внедрения соответствующих гарантий в целях уменьшения рисков возникновения сбоев или непреднамеренных последствий, в том числе вследствие предвзятости в данных, алгоритмах и других областях.
- В процессе разработки, внедрения и применения ИИ в военной области необходимо обеспечивать надлежащее участие человека, в частности принимать соответствующие меры для задействования человеческого суждения и обеспечения контроля за применением силы.
- Соответствующий персонал должен проходить подготовку, которая позволяла бы ему в достаточной мере понимать, разьяснять и отслеживать результаты, получаемые на выводе систем ИИ, применяемых в военной области, включая системы с поддержкой ИИ, и доверять таким результатам. Необходимо и далее прилагать усилия по повышению степени объяснимости и отслеживаемости результатов применения ИИ в военной области.

Республика Корея выступает в поддержку проведения обсуждений и диалогов по вопросам дальнейшей разработки мер по обеспечению ответственного применения ИИ в военной области, в том числе посредством разработки системы международного регулирования, введения строгих протоколов тестирования и оценки; внедрения комплексных процедур проверки, оценки и сертификации, создания надежных национальных механизмов надзора, обеспечения процессов непрерывного контроля, принятия комплексных программ обучения и проведения занятий, усиления кибербезопасности и установления четких систем подотчетности.

Принятие надежных мер контроля и обеспечения безопасности имеет решающее значение, чтобы не допустить приобретения и ненадлежащего использования безответственными субъектами потенциально опасных средств ИИ, предназначенных для применения в военной области, включая системы, поддерживающие ИИ.

Республика Корея выступает за разработку эффективных мер укрепления доверия и надлежащих мер уменьшения риска, а также за обмен информацией и проведение консультаций между государствами, посвященных передовому опыту и извлеченным урокам.

Республика Корея подчеркивает необходимость не допускать использования средств ИИ для содействия распространению оружия массового уничтожения государствами и негосударственными субъектами и особо отмечает, что средства ИИ не должны препятствовать усилиям в области контроля над вооружениями, разоружения и нераспространения. Крайне важно сохранять человеческий контроль и участие человека во всех действиях, имеющих решающее значение для информационного обеспечения и осуществления суверенных решений, касающихся применения ядерного оружия, без ущерба для достижения конечной цели, которая заключается в построении мира, свободного от ядерного оружия.

Разработка, внедрение и применение средств ИИ и систем с поддержкой ИИ в военной области должны осуществляться таким образом, чтобы поддерживать международный мир и безопасность и не препятствовать его обеспечению.

### **Будущее регулирование искусственного интеллекта в военной области**

В процессе определения концепции будущего регулирования ИИ в военной области крайне важно сформировать общее понимание технологии ИИ, ее возможностей и ограничений, и общее понимание возможных применений ИИ в военной области, а также их последствий для международного мира и безопасности.

Важное значение имеет и наращивание потенциала, особенно для развивающихся стран, с тем чтобы способствовать их полноценному участию в дискуссиях по вопросам управления и содействовать принятию ответственного подхода к разработке, внедрению и применению ИИ в военной области, а также достижению общего понимания этих вопросов. Достижению общего понимания может способствовать также обмен знаниями, передовым опытом и информацией об извлеченных уроках.

Вследствие стремительного развития ИИ механизмы регулирования должны отличаться достаточной гибкостью, чтобы не отставать от него. Кроме того, Республика Корея поддерживает сбалансированный подход, учитывающий как возможности, так и риски. Слишком большой упор на факторы риска или чрезмерно ограничительный подход в сфере регулирования могут препятствовать инновациям и заслонять потенциал ИИ для поддержки международного мира и безопасности. Будущая система регулирования не должна служить барьером для инноваций, а скорее поддерживать их и играть роль фактора, способствующего ответственному применению ИИ в военной сфере.

Поскольку международное сообщество находится на начальном этапе выработки понимания последствий применения ИИ в военной области для международного мира и безопасности и изучает текущее состояние развития технологий и политики, представляется преждевременным сужать направления регулирования ИИ или же принимать юридически обязывающие документы или нормы в отсутствие единого и общепринятого понимания концепции ответственного применения ИИ в военной области. Республика Корея считает, что обсуждение вопросов регулирования должно быть реалистичным и проходить постепенно, на основе постоянного диалога.

Признавая, что источником инноваций в сфере ИИ является частный сектор, Республика Корея считает, что будущие усилия по регулированию должны основываться на открытом и инклюзивном подходе, предусматривающем взаимодействие с многочисленными заинтересованными сторонами, включая

отраслевые структуры, научные круги, гражданское общество и региональные и международные организации.

Республика Корея отмечает национальные, региональные и глобальные усилия для использования возможностей и решения проблем в связи с применением ИИ в военной области, включая разработку соответствующих национальных стратегий, законодательства, принципов, норм, политики и мер, и признает важность поощрения диалога на всех уровнях.

Для обеспечения ответственного применения ИИ в военной области Республика Корея создала в Министерстве национальной обороны в 2022 году Отдел по политике в отношении данных и в 2025 году Группу по политике в отношении применения ИИ для целей обороны. В 2024 году Министерство создало Комитет по данным и ИИ для целей обороны в качестве высшего совещательного и директивного органа.

В сентябре 2024 года в Сеуле в целях поощрения диалога Республика Корея совместно с Нидерландами, Сингапуром, Кенией и Соединенным Королевством провела второй саммит по ответственному использованию искусственного интеллекта в военной сфере. Саммиты по ответственному использованию искусственного интеллекта в военной сфере и серия региональных консультаций по ответственному ИИ в военной области, прошедшая в 2024 году, послужили инкубатором для обмена опытом, развития всестороннего диалога и укрепления взаимопонимания. В перспективе третий саммит по ответственному использованию искусственного интеллекта в военной сфере, который пройдет в Испании в сентябре 2025 года, а также предстоящие региональные консультации по вопросам ответственного ИИ в военной области, которые пройдут в 2025 году, будут и далее направлять усилия международного сообщества по ответственному применению ИИ в военной области.

Республика Корея считает, что обсуждения вопросов ответственного применения ИИ в военной области в рамках Организации Объединенных Наций, включая Первый комитет Генеральной Ассамблеи и Комиссию по разоружению, должны дополнять другие соответствующие инициативы за пределами Организации Объединенных Наций, включая процесс проведения саммитов по ответственному использованию искусственного интеллекта в военной сфере, Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях и Группу правительственных экспертов по новым технологиям в сфере создания смертоносных автономных систем вооружений. Республика Корея считает, что эти инициативы носят взаимосоупротивляющий и взаимодополняющий характер.

Управление данными также имеет решающее значение. Поскольку данные играют центральную роль в обучении операторов, внедрении и оценке систем ИИ, соответствующие заинтересованные стороны должны продолжить обсуждение надлежащих механизмов управления данными, включая четкую политику и процедуры в вопросах сбора, хранения, обработки, обмена и удаления данных, а также обеспечения их защиты.

## Российская Федерация

[Подлинный текст на русском языке]  
[10 апреля 2025 года]

Российская Федерация отмечает принятие резолюции [79/239](#) Генеральной Ассамблеи от 24 декабря 2024 года и в соответствии с ее пунктом 7 имеет честь представить свой национальный вклад к докладу Генерального секретаря Организации Объединенных Наций восьмидесятой сессии Генеральной Ассамблеи 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 по негуманному оружию по смертоносным автономным системам вооружений в качестве оптимальной площадки для такой дискуссии. Именно профильная Группа правительственных экспертов 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 смертоносных автономных систем вооружений 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. 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 Организации Объединенных Наций по разоружению 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:

- a) содержать описание типов систем вооружений и военной техники с использованием ИИ, отдельные важные особенности их применения;
- b) не ограничиваться текущим пониманием указанных средств, а учитывать возможности их перспективного развития;
- c) быть универсальным с точки зрения понимания экспертным сообществом, включающим ученых, инженеров, техников, военнослужащих, юристов и специалистов по этике;
- d) не толковаться как ограничивающее технический прогресс и наносящее ущерб исследованиям в области мирной робототехники и ИИ;
- e) 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.

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

Общепризнано, что существующее международное право, в том числе международное гуманитарное право, 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 международное гуманитарное право, 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 международное гуманитарное право), 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 международное гуманитарное право, в ситуациях вооруженных конфликтов остается одним из приоритетов Российской Федерации. Вооруженные Силы Российской Федерации строго руководствуются закрепленными в федеральных и ведомственных правовых актах нормами международного гуманитарного права. В распорядительных документах и учебных программах для всех категорий военнослужащих отражаются вопросы соблюдения международного гуманитарного права, в том числе касающиеся применения новых видов вооружений. В 2022 году принята «Концепция деятельности Вооруженных Сил Российской Федерации в сфере разработки и применения систем вооружений с использованием технологий искусственного интеллекта».

Российское законодательство в полном объеме учитывает руководящие принципы в отношении систем вооружений с использованием технологии ИИ, одобренные консенсусом государствами — участниками Конвенции по негуманному оружию в 2019 году. 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 путем прямого управления.

Контроль может осуществляться за счет:

- a) повышения надежности и отказоустойчивости системы;
- b) ограничения типов целей;
- c) ограничения по продолжительности функционирования, географическому охвату и масштабу применения;
- d) своевременного вмешательства и деактивации;
- e) проведения испытаний систем вооружений и военной техники с использованием технологии ИИ в реалистичных оперативных условиях;
- f) допуска к управлению (контролю) лиц, успешно освоивших процедуры использования средств с применением ИИ;
- g) контроля процесса производства отдельных элементов и изделия в целом;
- h) 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 вовлеченности человека», «соответствующих контексту человеческого контроля и оценки», «предсказуемости, надежности, 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 международное гуманитарное право, 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 развертывания новых технологий в любых системах военного назначения.

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

- a) эффективного управления жизненным циклом;
- b) организации разносторонних испытаний на всех этапах жизненного цикла, в том числе в условиях, близких к реальной обстановке;
- c) обеспечения надежности и отказоустойчивости;
- d) определения критериев готовности;
- e) обеспечения максимальной защищенности от несанкционированного доступа;
- f) подготовки операторов;

g) приоритета в использовании технологии ИИ в процессе сбора и обработки информации, необходимой для поддержки принятия решений в военной сфере;

h) непрерывного контроля за действиями таких систем со стороны оператора, обеспечение возможности экстренного прекращения выполнения боевой задачи по его команде;

i) 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 Группе правительственных экспертов по смертоносным автономным системам вооружений как оптимальной международной площадке для такой дискуссии, в рамках действующих режимов в области контроля над вооружениями, разоружения и нераспространения, а также Комиссии по разоружению. При этом соответствующее обсуждение на одной площадке не должно дублировать обмен мнениями, который уже ведется 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 ее рассмотрения или ее обсуждение в узком составе, без участия подавляющего большинства государств — членов Организации Объединенных Наций (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 международное гуманитарное право, 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 [79/239](#) Генеральной Ассамблеи, а также включить настоящий документ в приложение к Вашему докладу.

## Сербия

[Подлинный текст на английском языке]  
[4 апреля 2025 года]

Развитие и применение искусственного интеллекта является одним из важных факторов, меняющих способы ведения военных операций в современном мире. Искусственный интеллект открывает новые возможности и в то же время порождает новые проблемы для международной стабильности, мира и безопасности в военной области. Поэтому необходимо приступить к созданию соответствующей международной системы, регулирующей его применение.

### 1. Возможности и преимущества, обеспечиваемые применением искусственного интеллекта в военной области

Несмертоносное военное применение искусственного интеллекта может улучшить многие аспекты военных операций:

- a) повысить уровень оперативной информированности;
- b) улучшить процесс принятия решений с точки зрения качества и скорости;
- c) повысить качество разведывательной информации и рекогносцировки путем оперативной обработки данных и поддержки оперативного выявления угроз;
- d) поддержать защиту гражданского населения и некомбатантов в военных конфликтах;
- e) поддержать операции и миссии в пользу мира путем мониторинга прекращения огня и прогнозирования динамики конфликта;
- f) усовершенствовать процессы и процедуры профилактического технического обслуживания и оптимизировать логистику за счет снижения затрат и экономии ресурсов.

### 2. Основные проблемы и угрозы, связанные с применением искусственного интеллекта в военной области

Развитие и внедрение искусственного интеллекта в боевые и небоевые системы создают серьезнейшую проблему для международного мира и стабильности, а также для международного гуманитарного права, прежде всего в следующих аспектах:

- a) технические риски и сбои в работе, вызванные ошибками приложений, которые проявляются в динамичной среде и могут угрожать жизни людей, наносить материальный ущерб и влиять на соблюдение норм международного гуманитарного права;
- b) правовые и этические риски, связанные с соблюдением норм международного права, особенно таких его принципов, как избирательность, соразмерность и принятие мер предосторожности при выборе целей;

- c) отсутствие четких правил, устанавливающих ответственность за акты и действия, совершаемые под управлением искусственного интеллекта;
- d) несовершенство алгоритмов, которое, предположительно, может становиться причиной предвзятости, ошибок в процессе принятия решений и дискриминации, поскольку применение нерепрезентативных групп данных может давать ошибки в идентификации гражданских лиц или создавать угрозу для этнических или национальных групп;
- e) применение алгоритмов искусственного интеллекта, которое может создать ложное впечатление о понижении уровня ответственности лиц, участвующих в процессе проведения операций;
- f) стратегические риски в процессе принятия решений с применением искусственного интеллекта на основе ошибочных предположений;
- g) неизбирательное сближение и стыковка с новыми технологиями, особенно в таких сферах, как проведение информационных и киберопераций или применение ядерных, химических и биологических средств;
- h) нехватка профессиональных кадров для разработки, организации и ответственного применения систем искусственного интеллекта в конфликтах;
- i) ненадлежащее использование искусственного интеллекта в рамках информационных операций путем создания и распространения дезинформации, которая может спровоцировать конфликты и обострить напряженность.

### 3. Создание правовых и этических рамок

Принимая во внимание проанализированные риски и проблемы, необходимо создать в рамках международного сообщества обязательные правовые и этические рамки в следующих целях:

- a) поощрение диалога в рамках Организации Объединенных Наций и принятие мер для начала его проведения с целью активизации соблюдения норм международного гуманитарного права, включая принятие международно-правовых норм, правил и принципов, которые обеспечивали бы соответствие разработки и применения систем искусственного интеллекта принципам международного гуманитарного права (разграничение, пропорциональность и меры предосторожности для защиты лиц, не участвующих в вооруженных конфликтах);
- b) инициирование процесса оценки законности применения систем и вооружений в свете разрешенных видов применения искусственного интеллекта;
- c) обеспечение защиты жизни и свободы людей во время вооруженных конфликтов и их частной жизни в мирное время, в частности, в контексте мониторинга;
- d) усиление механизмов Организации Объединенных Наций путем обеспечения обязательного изучения риска применения искусственного интеллекта в военных целях, модернизации Конференции по разоружению, согласования работы Комиссии по разоружению, создания новых специализированных органов Организации Объединенных Наций и расширения существующих инициатив Организации Объединенных Наций по ответственному применению искусственного интеллекта;

e) инициирование диалога в рамках Организации Объединенных Наций в целях выработки определения ответственного применения искусственного интеллекта в военной области и разработки протоколов безопасности для его применения (тестирование, оценка, проверка и сертификация);

f) разработка мер по обеспечению учета принципов международного гуманитарного права при привлечении частного сектора к разработке, создании и применении систем и услуг искусственного интеллекта в военных областях;

g) расширение существующих институтов и документов Организации Объединенных Наций с учетом рекомендаций, касающихся этики разработки и применения искусственного интеллекта, с включением в них конкретных руководящих принципов в отношении ведения конфликтов.

Применение систем искусственного интеллекта в контексте международных вооруженных конфликтов требует принятия широких многосторонних мер со стороны международного сообщества с целью повышения степени ответственности за их применение. Организация Объединенных Наций должна играть ведущую роль в налаживании диалога, разработке норм и наращивании потенциала международного сообщества во избежание распыления усилий и в целях обеспечения надлежащего управления.

## Сингапур

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Будучи малым государством, Сингапур неизменно выступает в поддержку многосторонней системы, основанной на правилах, и роли Организации Объединенных Наций. Организация Объединенных Наций обеспечивает основу для международного права и норм. Многосторонние структуры, системы и правовые положения имеют принципиально важное значение для выживания всех государств, особенно малых.

Сингапур считает, что необходимо обеспечить ответственный подход к применению средств ИИ в военной области, включая системы с поддержкой ИИ, на протяжении всего их жизненного цикла и их соответствие применимым нормам международного права, в частности международного гуманитарного права.

ИИ способен обеспечить преимущества в военной области путем повышения точности средств поражения и улучшения ситуационной осведомленности, и, соответственно, ограничения сопутствующего ущерба гражданскому населению и/или гражданским объектам. Вместе с тем в отсутствие надлежащих механизмов регулирования ИИ может также порождать риски эскалации конфликтов и возникновения просчетов. В этой связи Сингапур считает важным, чтобы международное сообщество занялось обсуждением этой темы.

### **Подход Сингапура к вопросам регулирования искусственного интеллекта в военной области**

Одна из ключевых целей версии 2.0 Национальной стратегии Сингапура в области искусственного интеллекта является создание среды, позволяющей обеспечивать надежную защиту пользователей и способствующей инновациям. С этой целью различные государственные структуры, включая оборонный сектор, разрабатывают рамочные основы регулирования ИИ, позволяющие использовать преимущества ИИ и обеспечивающие уменьшение потенциального вреда от его применения.

По итогам консультаций со специалистами по оборонным технологиям, военному планированию, международному праву и политике Сингапур разработал обнародованные в 2021 году национальные принципы, регулирующие применение ИИ в военной области, которые касаются четырех ключевых проблемных областей:

а) **ответственность.** Во-первых, необходимо устранить риски, связанные с поведением развивающегося ИИ. Системы ИИ должны иметь четко определенное целевое назначение, а ответственность за результаты работы систем ИИ несут как разработчики, так и пользователи;

б) **отказоустойчивость.** Во-вторых, необходимо устранить риски возникновения ошибок или неточностей в результатах работы системы ИИ. Системы ИИ должны проходить тестирование, а их надлежащее функционирование должно быть гарантировано на уровне, соответствующем их предполагаемому применению. На этапе их разработки необходимо сводить к минимуму вероятность непреднамеренной предвзятости и обеспечивать последовательность получаемых результатов;

в) **надежность.** В-третьих, необходимо устранить риски, связанные с использованием ИИ злоумышленниками. Системы ИИ должны разрабатываться с учетом кибернетических и иных угроз со стороны противника, связанных с ИИ. Во избежание возникновения «эффекта черного ящика» необходимо надлежащим образом документировать процесс разработки, чтобы обеспечивать объяснимость;

г) **безопасность.** В-четвертых, мы должны уделять особое внимание предотвращению риска возникновения сбоев в работе ИИ в условиях действия особых требований к обеспечению безопасности. Системы ИИ должны обеспечивать безопасное использование не только на уровне базовой платформы, но и для сопредельных ресурсов и персонала.

Изложенные руководящие принципы легли в основу принятого Сингапуром подхода к управлению разработкой, тестированием, обучением и внедрением систем с поддержкой ИИ в военных целях.

#### **Международные и региональные инициативы по вопросам применения искусственного интеллекта в военной области**

Сингапур активно участвует в международных инициативах по вопросам регулирования применения ИИ в военной области. В 2023 году Сингапур поддержал Призыв к действиям по ответственному применению искусственного интеллекта в военной области и Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях. В 2024 году Сингапур выступил в качестве одного из организаторов саммита по ответственному использованию искусственного интеллекта в военной сфере в Сеуле, Республика Корея, на котором мы одобрили План действий по ответственному использованию искусственного интеллекта в военной сфере.

Кроме того, Сингапур признает важность региональных инициатив для обеспечения всестороннего и контекстуального обсуждения вопросов применения ИИ в военной области. Сингапур выступил в качестве одного из организаторов региональных консультаций по вопросам ответственного применения ИИ в военной области для Азии в 2024 году, которые стали платформой для обмена мнениями между странами региона, в том числе о возможностях и рисках, связанных с применением ИИ в военной области.



В феврале 2025 года Сингапур и другие государства — члены Ассоциации государств Юго-Восточной Азии (АСЕАН) приняли совместное заявление о сотрудничестве в сфере, касающейся применения ИИ в оборонном секторе, на выездном совещании министров обороны стран АСЕАН в Пенанге, Малайзия. В этом заявлении министры обороны стран АСЕАН обязались поощрять подотчетность и ответственность в сфере применения ИИ, углублять региональное понимание и осведомленность о последствиях применения ИИ в оборонном секторе путем обмена информацией, а также обмениваться передовым опытом и информацией об извлеченных уроках между государствами — членами АСЕАН.

### **Дальнейшие перспективы обсуждения вопросов искусственного интеллекта и международного мира и безопасности в Организации Объединенных Наций**

Сингапур считает, что любые дальнейшие обсуждения, направленные на укрепление поддержки этой резолюции со стороны международного сообщества, должны носить открытый и инклюзивный характер. В этой связи мы положительно отнеслись бы к идее создания в рамках Организации Объединенных Наций рабочей группы открытого состава, в центре внимания которой находилась бы проблема применения ИИ в военной области. Если такая рабочая группа открытого состава будет создана, то она должна будет принять многосторонний подход с привлечением, в частности, специалистов по технологиям, военному планированию, международному праву и политике. Мы подтверждаем нашу приверженность сотрудничеству со всеми государствами-членами в целях содействия ответственному применению ИИ в военной области.

### **Испания**

[Подлинный текст на испанском языке]  
[11 апреля 2025 года]

### **Введение**

Развитие искусственного интеллекта (ИИ) означает революцию во всех сферах, включая безопасность и оборону. Разработка и внедрение средств ИИ способствуют достижению значительного прогресса и открывают широкие возможности, но в то же время порождают многочисленные проблемы.

Внедрение этой технологии в военной области не только преобразует военные операции, но и изменит общий стратегический баланс.

Разработка и внедрение ИИ в Министерстве обороны основываются на принципах ответственного, этичного и законного применения в военной области в соответствии с нормами международного гуманитарного права и с соблюдением прав человека.

ИИ меняет традиционное представление о военной мощи и безопасности, открывая более широкие возможности для сбора и анализа данных, принятия решений и проведения операций в разных средах. Это подразумевает изменение общепринятого подхода государств к обеспечению обороны и безопасности, способствуя более оперативному и точному реагированию на возникающие угрозы.

В военной области ИИ коренным образом меняет непредсказуемую ситуацию на театре военных действий, знаменуя собой пересмотр шаблонов планирования и проведения военных операций. Он затрагивает и другие военные аспекты: тыловое обеспечение, обучение личного состава, управление информацией и ее анализ, разведку, наблюдение, обнаружение целей и рекогносцировку.

Следует отметить, что Испания, демонстрируя свою приверженность ответственному использованию ИИ, выступила принимающей страной для Саммита по ответственному использованию искусственного интеллекта в военной сфере, состоявшегося в 2025 году, а также присоединилась к «Призыву к действиям» (Гаага, 2023 год) и «Плану действий», представленному на Саммите 2024 года.

### **Концептуальная и нормативная основы, принятые в Министерстве обороны**

Разработка и внедрение приложений ИИ в Министерстве обороны осуществляются с соблюдением ряда основополагающих принципов, которые обеспечивают их безопасное и этичное применение, соответствующее национальному законодательству и нормам международного права. Эти принципы закреплены в «Стратегии разработки, внедрения и применения искусственного интеллекта в Министерстве обороны», разработанной в соответствии с постановлением 11197/2023 Министерства обороны и согласованной со стратегией Организации Североатлантического договора (НАТО) в области ИИ 2021 года (пересмотрена в 2024 году), и имеют целью обеспечение максимально широкого использования возможностей ИИ в сфере обороны, а также уменьшение рисков, связанных с его применением в военной области, как об этом говорится ниже.

- **Законность:** разработка и применение приложений ИИ должны соответствовать применимому национальному законодательству и нормам международного права, включая Всеобщую декларацию прав человека и нормы международного гуманитарного права.
- **Ответственность и подотчетность человека:** любые разработки и виды применения ИИ должны предусматривать четкий контроль со стороны человека для обеспечения надлежащей подотчетности и ответственности.
- **Понятность и отслеживаемость:** приложения ИИ должны быть понятны и транспарантны для соответствующего персонала, включая использование проверяемых методологий, источников и процессов.
- **Надежность и транспарентность:** необходимо четко определить и ограничить случаи применения ИИ и предоставлять информацию для общего понимания приложений ИИ всеми заинтересованными сторонами. Необходимо проводить тестирование и обеспечивать гарантии безопасности, защищенности и отказоустойчивости этих приложений в рамках различных сценариев применения на протяжении всего жизненного цикла таких приложений.
- **Управление:** приложения ИИ необходимо разрабатывать и применять в соответствии с их предполагаемым назначением, обеспечивая возможность выявления и предотвращения непредвиденных последствий. В случае обнаружения незапланированного или нежелательного поведения необходимо задействовать механизмы отключения или деактивации.
- **Уменьшение предвзятости:** необходимо принимать все соответствующие меры для минимизации ошибок и субъективных решений в процессе разработки и применения средств ИИ.
- **Конфиденциальность:** разработка, внедрение и использование приложений с поддержкой ИИ должны обеспечивать неприкосновенность частной жизни людей с момента проектного задания и на протяжении всего жизненного цикла.

Что касается нормативно-правовой базы, то для разработки, внедрения и применения ИИ в военной области в Министерстве обороны разрабатывается комплекс стандартов и примеров передовой практики, обеспечивающих его ответственное и эффективное применение в соответствии с национальным законодательством и нормами международного права, в частности, при строгом соблюдении норм международного гуманитарного права и прав человека.

### **Возможности**

Министерство обороны уделяет особое внимание наращиванию потенциала ИИ в различных областях, чтобы повысить эффективность вооруженных сил. Согласно стратегии, ИИ применяется для целей ведения операций, разведки, материально-технического снабжения и обеспечения кибербезопасности, а также для поддержки принятия решений.

ИИ позволит повысить точность, скорость и эффективность принятия решений в ходе военной операции при полном соблюдении норм международного гуманитарного права, что повысит эффективность выполнения задач и уменьшит риск для личного состава, а также будет способствовать усилению защиты гражданского населения и гражданских объектов в вооруженных конфликтах.

Его способность анализировать большие объемы данных в режиме реального времени повышает ситуационную осведомленность и возможности реагирования на угрозы, что способствует укреплению оперативной безопасности. В процессе совершенствования возможностей всегда обеспечивается контроль со стороны человека и не допускается перекладывание ответственности на машины.

Что касается военной подготовки и обучения в рамках группы командно-штабных училищ пяти европейских стран (группа училищ Соединенного Королевства, Франции, Германии, Италии и Испании), то в них проводится работа по налаживанию сотрудничества по вопросам ИИ в сфере военного образования.

Кроме того, Испания сотрудничает с Советом НАТО по искусственному интеллекту и обработке данных в сфере ответственного использования данных и ИИ в военной области.

Министерство обороны объявило также о стратегических инвестициях в определенные сферы с целью стимулирования проектов, касающихся ИИ и других передовых технологий. Эти инвестиции призваны не только укрепить эту отрасль, но и содействовать оживлению промышленности новых региональных территорий.

### **Проблемы**

Разработка и применение ИИ в военной области должны соответствовать национальному законодательству и нормам международного права, включая нормы международного гуманитарного права, поэтому в настоящее время ведется работа по обеспечению значимого человеческого контроля за принятием важнейших решений, касающихся применения ИИ в военных операциях.

Что же касается конфиденциальности и защиты данных, то сбор и обработка больших объемов данных для обучения моделей ИИ порождают риски с точки зрения защиты персональных данных и обеспечения информационной безопасности.

*Безопасность и надежность*

Главная проблема заключается в обеспечении безопасности и надежности ИИ, а основные связанные с этим риски заключаются в следующем:

- алгоритмы искусственного интеллекта могут основываться на предвзято подобранных обучающих данных, что может становиться причиной ошибочных решений или непреднамеренных последствий;
- плохое обучение моделей ИИ может вызывать ошибки в интерпретации с катастрофическими последствиями в ходе проведения военных операций;
- системы ИИ могут подвергаться кибератакам с целью манипулирования их поведением или выведения их из строя;
- существует риск отравления данных, когда злоумышленники изменяют обучающие наборы данных, чтобы вызвать сбой в работе алгоритмов.

В Испании разработка ИИ в военной области регулируется принципами подотчетности и постоянного контроля на основе задействования механизмов оценки рисков, проверки и отслеживания на всех этапах жизненного цикла систем. Любые разработки или виды применения ИИ должны предусматривать четкий надзор со стороны человека, чтобы обеспечивать надлежащую подотчетность и ответственность, позволяя четко отслеживать действия человека, связанные с работой ИИ и осуществляемые параллельно с использованием ИИ, и не допуская принятия окончательных решений машинами.

Кроме того, ИИ должен обеспечивать надежность и предсказуемость, сохраняя при этом соответствующий уровень автономности под контролем и надзором соответствующим образом подготовленных операторов.

Любое средство ИИ должно оцениваться в среде, отличной от той, в которой оно обучалось, и должно проходить нефункциональное тестирование, например проверку под нагрузкой, в состоянии стресса и на работоспособность в определенных меняющихся сценариях, чтобы изучить его поведение и допустимые отклонения.

Кроме того, эти средства ИИ будут проходить тщательное тестирование и постоянную проверку на протяжении всего жизненного цикла, что позволит заблаговременно выявлять потенциальные ошибки и повышать их эксплуатационную надежность. На всех этапах развертывания будут применяться протоколы, обеспечивающие надзор и контроль со стороны человека и гарантирующие, что принятие важнейших решений не будет делегировано исключительно ИИ. В связи с этим проводится работа по сертификации создаваемых средств ИИ признанными структурами.

Для повышения степени надежности систем на основе ИИ и их защиты от внешнего воздействия важно еще на этапе проектной разработки заложить в них системы защиты, обеспечив их устойчивость к кибератакам и манипуляциям со стороны противника, а также целостность используемых данных и моделей.

ИИ может стать объектом таких атак, как отравление данных или манипулирование моделями, что требует постоянного мониторинга работы системы и проведения регулярных проверочных тестов и проверок. Следует поощрять разработку планов резервного копирования и послеаварийного восстановления, гарантирующих работоспособность систем в неблагоприятных сценариях.

Кроме того, рекомендуется поддерживать сотрудничество с агентствами по обеспечению кибербезопасности и специалистами по ИИ, что позволит вооруженным силам получить доступ к самым лучшим инструментам и стратегии для защиты таких систем от внешних угроз и обеспечения их эксплуатационной надежности.

Более того, талант и подготовка специалистов в области этих технологий имеют решающее значение и рассматриваются в качестве одного из четырех основных компонентов, которым Министерство обороны уделяет главное внимание, чтобы операторы понимали масштабы и ограничения этих систем и могли вмешиваться в случае отклонений в их поведении. Обучение операторов и их информирование о правовом и этическом применении ИИ имеют ключевое значение для уменьшения рисков, связанных с предвзятостью, чтобы применение ИИ в вооруженных силах было объективным и надежным и соответствовало национальному законодательству и нормам международного права, особенно международного гуманитарного права.

В настоящее время ведется работа над созданием руководства по передовому опыту, которое может стать основой для документа, подготовленного с участием всех структур Министерства обороны. Осуществляется распространение информации о предлагаемых передовых методах НАТО для ответственного применения ИИ в военной области. В качестве примера можно назвать инструментарий НАТО для оценки ответственного ИИ, цель которого заключается во внедрении принятых НАТО принципов ответственного применения ИИ, включая законность, подотчетность, отслеживаемость, надежность, управляемость и уменьшение предвзятости.

## Швейцария

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

### 1. Возможности и риски

Искусственный интеллект (ИИ), вероятно, изменит многие аспекты военного дела. В перспективе он будет способен оказывать поддержку в решении задач и ведении боевых действий, например, благодаря повышению уровня надежности, эффективности, точности, безопасности и безотказности. Ключевые сферы применения ИИ включают обеспечение ситуационной осведомленности, принятие решений, разведку, наблюдение и рекогносцировку, тыловое обеспечение и систему снабжения, обучение и моделирование, а также командование и управление путем анализа больших массивов данных и принятия более оперативных и обоснованных решений. Например, в сфере наблюдения и разведки ИИ может анализировать снимки с беспилотных летательных аппаратов и спутников, выявляя изменения быстрее, чем люди-специалисты по анализу. Кроме того, ИИ может поддерживать распознавание целей, обрабатывая данные, поступающие с датчиков, чтобы определять свои силы и силы противника. В сфере тылового обеспечения ИИ может оптимизировать систему снабжения, предсказывать поломки материальной части и обеспечивать доставку ресурсов в нужное место и в нужное время. В вопросах поддержки принятия решений моделирование с применением ИИ позволит командирам получать аналитические оценки и прогнозируемые результаты для целей стратегического планирования. Системы обучения и моделирования на базе ИИ обеспечивают реалистичную и адаптивную среду, которая позволит повысить уровень подготовки личного состава. Наконец, ИИ может обеспечивать поддержку функции командования и управления путем оптимизации потока информации,

совершенствования процесса принятия решений и повышения уровня координации между подразделениями. ИИ может также являться подспорьем в выявлении угроз, обеспечении кибербезопасности, деятельности по поддержанию мира, проверке контроля над вооружениями и деэскалации конфликтов благодаря использованию систем раннего предупреждения, аналитических оценок и механизмов мониторинга, способствуя укреплению стабильности и безопасности. Вместе с тем, хотя эти разработки могут принести пользу вооруженным силам, внедрение ИИ в военной области порождает также ряд серьезных опасений и возможных рисков.

При условии ответственного применения в вооруженном конфликте ИИ может способствовать обеспечению соблюдения норм международного гуманитарного права и усилению защиты гражданского населения и гражданских объектов, например путем совершенствования оценки рисков или повышения точности целеуказания для уменьшения сопутствующего ущерба. Однако некоторые виды применения ИИ в военной области в условиях вооруженного конфликта, особенно приложения, связанные с высоким риском, также порождают требующие решения серьезные проблемы правового, гуманитарного и этического характера, а также проблемы безопасности и стратегической стабильности, например:

- **Ошибки в выборе цели.** Технически ИИ способен идентифицировать объекты или людей на основе данных, используемых для его обучения, однако обеспечение понимания обстановки и применения ценностных суждений, необходимых для соблюдения норм международного права, представляет собой особую проблему, которая могла бы привести к ошибочной идентификации объектов или людей в качестве военных целей и, следовательно, стать причиной незаконных или непреднамеренных ударов.
- **Риски эскалации.** В условиях стремительно развивающегося кризиса средство поддержки принятия решений, функционирующее по принципу «черного ящика», могло бы рекомендовать агрессивное воздействие без четкого обоснования. В отсутствие объяснимости командиры могут либо слепо следовать ошибочным указаниям, либо тратить критически важное время на их оспаривание.
- **Неверное толкование намерений.** Применение системы ИИ для оценки риска, связанного с действиями людей и/или с объектами, может вызывать опасения (с точки зрения права и безопасности), особенно если оценки основываются на шаблонах, определенных на основе прошлого поведения и условий, без контекстуального контроля со стороны человека и его участия в принятии решений. Например, система ИИ, отслеживающая поведение противника, может вследствие получения некачественных данных ошибочно классифицировать обычные перемещения войск как враждебные действия, что может привести к нанесению превентивного удара и вызвать непреднамеренную эскалацию.

Эти риски подчеркивают не только обязанность обеспечивать соблюдение действующих норм международного права, в частности международного гуманитарного права, но и настоятельную необходимость дальнейшего диалога и дальнейшего изучения этого вопроса ради лучшего понимания рисков и проблем, возможных необходимых мер, а также для изучения необходимости, дополнительной пользы и целесообразности разработки дополнительных структур нормативного регулирования. Это могло бы включать разработку национальных законов, обобщение передового опыта, выработку международных норм, стандартов или документов, а также принятие практических рекомендаций.

## 2. Правовая основа

Разработка и применение ИИ, как и любой другой технологии, не происходят в правовом вакууме. Разработка, внедрение и применение ИИ в военной области должны осуществляться в полном соответствии с действующими нормами международного права, в частности с Уставом Организации Объединенных Наций, нормами международного гуманитарного права и права прав человека, а также другими соответствующими правовыми основами. Никакие технологии никогда не должны ставить под сомнение действие норм международного права. Нормы международного права, в частности Устав Организации Объединенных Наций во всей его полноте, нормы международного права прав человека и международного гуманитарного права, сохраняют свою применимость и подлежат соблюдению и выполнению.

Государства и стороны в конфликте должны при любых обстоятельствах соблюдать и обеспечивать соблюдение норм международного гуманитарного права, в частности в процессе применения ИИ в ходе военных действий. Таким образом, разработки в сфере военного ИИ должны способствовать обеспечению соблюдения норм международного гуманитарного права и защиты гражданского населения и гражданских объектов. Этого можно достичь, например, запрограммировав системы ИИ на уделение главного внимания обеспечению точности, сведению вреда к минимуму и обеспечению подотчетности, например, с помощью процессов строгого выбора, проверки и подтверждения целей. Кроме того, в процессе применения ИИ следует повысить эффективность выполнения правового обязательства по принятию всех возможных мер предосторожности в ходе ведения боевых действий, включая предотвращение или, по крайней мере, сведение к минимуму случайного ущерба, путем оказания командирам поддержки в защите гражданского населения и гражданских объектов в ходе боевых действий, например, путем совершенствования оценки рисков.

Одно из ключевых направлений заключается в разработке и обучении военного ИИ на наборах данных, позволяющих использовать его в полном соответствии с нормами международного права. За рамками боевых действий применение ИИ в военной области должно соответствовать всем соответствующим нормам и принципам международного гуманитарного права, если ИИ используется для решения других задач, регулируемых нормами международного гуманитарного права, например, в связи с задержанием и интернированием людей или в связи с борьбой с беспорядками и принятием мер по обеспечению общественной безопасности на оккупированных территориях.

В процессе разработки и применения ИИ в военной области существует риск того, что в проект системы или в данные, используемые для ее обучения, могут быть заложены чрезмерно либеральные юридические трактовки, например, расширенное определение законных целей или завышенные пороги допустимого случайного ущерба. При масштабном применении такие трактовки могут постепенно подорвать защитную функцию норм международного гуманитарного права и значительно расширить ущерб гражданскому населению. Этот риск подчеркивает важность сохранения целостности правовых норм, которым необходимо уделять основное внимание в процессе регулирования, разработки и внедрения ИИ в военной области в будущем.

## 3. Договоренности и принципы

С учетом изложенных выше правовых рамок и в порядке их развития, а также принимая во внимание гуманитарные и этические соображения, соображения безопасности и стратегической стабильности, следует более детально проработать следующие договоренности и принципы:

1. *Ответственность, подотчетность и участие*

- **Ответственность и подотчетность.** Государства должны обеспечить, чтобы в соответствии с применимыми нормами международного права ответственность и подотчетность за принятие решений, касающихся применения ИИ в военной области, всегда лежали на человеке.
- **Контекстуальный контроль со стороны человека и его участие в принятии решений.** Критически важные военные решения на всех уровнях от ситуационного центра до театра военных действий, особенно решения, касающиеся применения силы, всегда должны находиться под контекстуальным контролем человека и приниматься с участием человека. Применение ИИ в военной области может помочь в принятии решений, но ИИ не должен подменять юридические и этические соображения и суждения, такие как когнитивная автономия при принятии решений. Государства должны интегрировать такие системы в структуру командования и управления только в тех случаях, когда принятие решений и надлежащий уровень контроля остаются за человеком. Следует в максимально возможной степени устранять непреднамеренную предвзятость.

2. *Надежность, предсказуемость/объяснимость, безотказность*

- **Надежность.** Средства ИИ, применяемые в военной области, должны отличаться надежностью, чтобы не допускать непреднамеренных последствий или сбоев, особенно если они могут оказать негативное воздействие на гражданское население и гражданские объекты или причинить им вред. ИИ в военной области должен применяться только в том случае, если можно разумно предвидеть его последствия.
- **Предсказуемость/объяснимость.** Процессы принятия решений системами ИИ должны быть предсказуемыми и объяснимыми для тех лиц, которые отвечают за их внедрение, что позволит им понимать и предвидеть поведение системы.
- **Безотказность.** Средства ИИ, применяемые в военной области, должны также отличаться безотказностью как в техническом, так и в эксплуатационном плане, чтобы сохранять безопасность и надежность в процессе внедрения и применения.

3. *Уменьшение риска*

- **Повышение уровня ситуационной осведомленности.** ИИ следует использовать для повышения степени осведомленности о ситуации на театре военных действий, в частности, для обнаружения присутствия гражданского населения с целью уменьшения вероятности причинения вреда.
- **Аналитические оценки.** Аналитические модели на основе ИИ должны использоваться для оценки рисков и для разработки, в частности, стратегий деэскалации конфликта и предотвращения жертв среди гражданского населения.
- **Встроенные ограничения.** Средства ИИ, применяемые в военной области, должны предусматривать защитные механизмы, которые сводят к минимуму вред и позволяют человеку осуществлять необходимое вмешательство в случае сбоя в работе системы.



#### 4. Недопущение новых вариантов эскалации

- **Стабильность.** Разработка, внедрение и применение средств ИИ в военной области должны осуществляться только таким образом, чтобы не усугублять международную напряженность и не создавать новых вариантов эскалации.
- **Контроль над вооружениями.** ИИ может поддерживать контроль над вооружениями и не должен подрывать существующие нормы и документы по вопросам нераспространения, контроля над вооружениями и разоружения или препятствовать соблюдению таких норм, особенно в отношении биологического и ядерного оружия.
- **Урегулирование кризисов.** Средства ИИ, применяемые в военной области, могут способствовать деэскалации и урегулированию кризисов.

#### 5. Управление жизненным циклом военных систем искусственного интеллекта

Ответственное военной применение ИИ требует принятия комплексного и учитывающего риски подхода, охватывающего весь жизненный цикл средств ИИ, применяемых в военной области. Это включает проектирование, разработку, тестирование, внедрение, эксплуатацию, обновление и вывод из эксплуатации таких средств. На всех этих этапах необходимо систематически учитывать соответствующие правовые, гуманитарные, оперативные и технические соображения. Такой подход, касающийся всего жизненного цикла, особенно важен для применяемых в военной области средств ИИ с высокой степенью риска, например, связанных с применением автономных вооружений, выбором целей или поддержкой принятия решений, способных причинить вред людям или вызвать гибель людей или повреждения объектов, и в целом в тех случаях, когда решения регулируются нормами международного гуманитарного права. Для систем с меньшей степенью риска, таких как средства административной поддержки или системы планирования тылового обеспечения, управление жизненным циклом должно осуществляться на основе оценки рисков в зависимости от контекста.

- На этапе проектирования и разработки государства должны обеспечивать обучение таких систем на высококачественных, репрезентативных наборах данных, содержащих минимум предвзятости, что позволит использовать такие системы в полном соответствии с нормами международного права и другими нормами и стандартами, максимально ограничивая нежелательные предубеждения.
- На этапе тестирования и оценки необходимо внедрить строгие процедуры проверки и сертификации, подтверждающие надежность, соответствие законодательству и безотказность эксплуатации в реальных условиях.
- На этапе внедрения и оперативного применения необходимо предусмотреть меры предосторожности для мониторинга работы системы, обеспечения контроля и принятия решений человеком с учетом контекста, а также для надлежащего вмешательства человека.
- На всех этапах обновления и обучения государства должны установить строгие протоколы для внесения изменений в систему, включая управление версиями, повторную сертификацию и официальные процессы утверждения.
- На этапе списания или вывода из эксплуатации необходимо принять меры для надежного отключения или архивирования систем, чтобы не допустить

их неправильного применения, непреднамеренной активации или повторного развертывания.

#### 4. Международное регулирование

Швейцария подчеркивает важность проведения инклюзивного и последовательного процесса в рамках Организации Объединенных Наций, направленного на укрепление общего понимания преимуществ, рисков и проблем, связанных с применением ИИ в военной области, и разработку принципов его ответственного применения. Поэтому к этому процессу необходимо привлечь все государства-члены и соответствующие заинтересованные стороны, а также ученых и представителей технологических отраслей, гражданского общества и академических кругов, чтобы обеспечить легитимность, специальные знания и широкую поддержку. Соответствующие процессы в рамках Организации Объединенных Наций должны носить транспарентный характер, регулярно созываться и согласовываться с другими соответствующими инициативами.

Главной целью всех усилий по международному регулированию ответственного применения ИИ в военной области должно быть обеспечение соблюдения норм международного права, в частности международного гуманитарного права. Кроме того, в центре таких усилий должны быть гуманитарные и этические соображения, обеспечение стабильности и уменьшение рисков безопасности. Эффективные механизмы регулирования, общие нормы и постоянный многосторонний диалог должны способствовать предотвращению непреднамеренной эскалации, обеспечению транспарентности и взаимного доверия, а также усиливать роль международного права в период технологических преобразований. Закладывая такие принципы в основу регулирования применения ИИ в военной области, государства способствуют созданию более предсказуемой, устойчивой и мирной обстановки в плане безопасности.

Конкретные усилия могли бы включать:

- содействие выработке общего понимания, определений и терминологии, а также определению общей сферы применения ИИ в военной области;
- выявление и углубление понимания возможностей и проблем, связанных с гуманитарным аспектом, законодательством, безопасностью и этикой;
- изучение мер по обеспечению транспарентности и укреплению доверия;
- разработка принципов, норм, подборок передовой практики и различных рекомендаций;
- принятие руководящих принципов по их внедрению.

#### Украина

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Украина активно разрабатывает и применяет искусственный интеллект (ИИ) в различных сферах деятельности, в том числе и в военной области. Украина четко осознает как потенциал этой технологии для повышения уровня благосостояния людей и наращивания военного потенциала, так и значительные риски ее неправомерного использования в гражданской и особенно в военной области. Эти риски особенно усиливаются в контексте неспровоцированного и неоправданного полномасштабного вторжения Российской Федерации на Украину, в ходе которого она систематически нарушает законы и обычаи войны и нормы международного гуманитарного права.

Украина поддерживает международные усилия по формированию глобального консенсуса в отношении ответственной разработки, внедрения и применения гражданского и военного ИИ и принимает участие в таких усилиях.

На сегодняшний день Украина, в частности, подписала Блетчлийскую декларацию в 2023 году; является одним из государств, поддержавших Политическую декларацию об ответственном применении искусственного интеллекта и автономных систем в военных целях, принятую на саммите по ответственному использованию искусственного интеллекта в военной сфере, который состоялся в Гааге в 2023 году; поддержала Призыв к действиям по ответственному применению искусственного интеллекта в военной области, согласованный на саммите по ответственному использованию искусственного интеллекта в военной сфере в 2023 году, и План действий по ответственному использованию искусственного интеллекта в военной области, принятый в качестве итогового документа саммита по ответственному использованию искусственного интеллекта в военной сфере в 2024 году; присоединились к Заявлению об инклюзивном и устойчивом использовании искусственного интеллекта для людей и планеты на саммите действий по искусственному интеллекту, прошедшем в 2025 года в Париже; и стал одним из авторов всех трех принятых на сегодняшний день резолюций Генеральной Ассамблеи по ИИ, включая резолюцию 79/239 о применении искусственного интеллекта в военной области и его последствиях для международного мира и безопасности.

Украина готова принять активное участие в новых глобальных инициативах, направленных на поощрение безопасного, этичного и ответственного развития ИИ. Она поддерживает также обсуждение вопроса об ИИ в его различных аспектах в рамках всей системы Организации Объединенных Наций, в том числе в Совете Безопасности.

Являясь одновременно миролюбивым государством, не имеющим территориальных претензий к другим, и жертвой российской военной агрессии, не признающей подобных претензий к себе, Украина разрабатывает и применяет военный ИИ исключительно для укрепления своей обороноспособности, осуществляя право на самооборону, предусмотренное Уставом Организации Объединенных Наций.

В сфере применения ИИ в военном контексте Украина выделяет следующие ключевые риски для международного мира и безопасности:

- Соперничество в сфере внедрения ИИ в боевые и оружейные системы, в особенности появление полностью автономных систем вооружений, действующих без вмешательства человека, может спровоцировать новый, более опасный виток глобальной гонки вооружений в ущерб достижению целей в области устойчивого развития.
- Как и другие цифровые технологии, с ростом угрозы кибератак, усложнением и расширением сфер применения ИИ, применяемый в военных системах, становится все более уязвимым для кибервмешательства и манипуляций со стороны заинтересованных субъектов в целях лишения таких систем их целевых характеристик и функции избирательного применения.
- Чрезмерная зависимость от ИИ в принятии решений может привести к потере человеческого контроля над критически важными военными процессами.

- Поспешное внедрение недоработанного ИИ в системы вооружений, особенно с несовершенными возможностями выявления целей, может привести к неизбирательному воздействию и увеличению числа жертв среди гражданского населения.
- В настоящее время не существует многосторонних рамок для контроля за распространением вооружений с встроенным ИИ.
- Применение вооружений с встроенным ИИ без соблюдения законов и обычаев войны и норм международного гуманитарного права вызывает серьезные правовые и этические проблемы.

## **Соединенное Королевство Великобритании и Северной Ирландии**

[Подлинный текст на английском языке]

[11 апреля 2025 года]

Искусственный интеллект (ИИ) представляет собой семейство технологий общего назначения, из которых любая может наделить машины способностью выполнять функции, традиционно требовавшие участия человеческого или биологического интеллекта, особенно если машины учатся выполнять эти функции на основе данных. Развитие и внедрение технологий ИИ происходит необычайно быстрыми темпами. В качестве группы технологий с различными системами, методами и приложениями, они имеют разные траектории развития, а их применение порождает разные последствия. Нет никаких сомнений в том, что они способны вызвать революционные преобразования во всех аспектах жизни общества, экономики и политики, включая оборону и безопасность.

Соединенное Королевство приветствует предоставленную резолюцией [79/239](#) Генеральной Ассамблеи возможность изучить последствия применения ИИ в военной области помимо смертоносных автономных систем вооружений в контексте широких и ценных обсуждений этого вопроса, в том числе в рамках Группы правительственных экспертов, созданной в соответствии с Конвенцией о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие. Тщательная оценка более широких стратегических последствий применения военного ИИ на основе всей совокупности мыслей, идей и передового опыта, обсуждаемых на неофициальных и официальных международных форумах по этой повестке дня, позволит провести целостную дискуссию по вопросу о способах максимально широкого использования возможностей ИИ в военной области наряду с принятием эффективных мер по устранению связанных с ним рисков.

### **Возможности, связанные с применением искусственного интеллекта в военной области**

Внедрение ИИ в военной области способно преобразовать оборону, динамику в сфере обеспечения глобальной безопасности и характер военных действий. Передовые технологии с поддержкой ИИ, которые позволяют быстрее и полнее классифицировать и уточнять большие объемы данных из разных источников, будут способствовать повышению эффективности и совершенствованию процесса принятия решений, а также повышению скорости и точности оперативного планирования. Применение ИИ в системах разведки, наблюдения и рекогносцировки может обеспечить более точное представление об оперативной обстановке и позволить органам планирования уменьшить последствия для гражданского населения, что приведет к повышению уровня защиты

гражданского населения и гражданской инфраструктуры. Автоматизация функций тылового обеспечения и поиска неразорвавшихся боеприпасов уменьшит необходимость присутствия личного состава на местности. Таким образом, применение ИИ в военном контексте может укрепить национальную и международную безопасность, уменьшить риск для жизни людей и сократить количество жертв.

Проведенное Министерством обороны Соединенного Королевства исследование по вопросам ИИ и поддержания мира позволило определить, каким образом можно использовать средства и системы с поддержкой ИИ во благо операций по поддержанию мира, в частности:

- задействование аналитического потенциала в целях повышения уровня ситуационной осведомленности, совершенствования процесса принятия оперативных решений, планирования с учетом вариантов обстановки и потенциала для анализа настроений;
- применение таких автономных систем, как беспилотные летательные аппараты, в целях расширения охвата больших географических районов или регионов повышенного риска (где постоянное присутствие миротворческого персонала может быть сопряжено с риском);
- материально-техническое снабжение в целях улучшения ситуации с оказанием медицинской и иной помощи местному населению в поддержку целей миссии и укрепления доверия в обществе.

Указанные возможности могут использоваться для усиления мониторинга и проверки соблюдения соглашений о контроле над вооружениями и мирных соглашений, что облегчит своевременное и достоверное выявление нарушений или подтверждение соблюдения. Средства ИИ могут обеспечивать более надежное обнаружение, идентификацию, атрибуцию и проверку враждебных подпольных операций различного рода, что позволит понизить эффективность таких операций а, возможно, и полностью предотвратить их. Кроме того, они могут способствовать отслеживанию и выявлению в режиме реального времени ненавистнических высказываний, пропаганды или изменений в общественных настроениях, которые могут привести к эскалации напряженности или подорвать любые мирные переговоры или режим прекращения огня.

### **Проблемы и риски**

Применение ИИ в военном контексте может усугубить существующие риски и создать дополнительные угрозы, вызывающие вооруженный конфликт или подводящие к порогу его развязывания. Поспешное внедрение средств ИИ ради получения стратегических преимуществ может привести к тому, что страны будут использовать ИИ неприемлемым образом с точек зрения права, этики или безопасности. Новые риски эскалации или аварий, вызванных неисправностями или же уязвимостью, хрупкостью, незрелостью или ненадежностью систем ИИ, потребуют принятия новых протоколов и механизмов деэскалации. Враждебно настроенные субъекты могут попытаться атаковать национальные системы ИИ и вызвать сомнения в их работоспособности, безопасности и надежности (например, путем «отравления» источников данных, порчи аппаратных средств в системах поставок и вмешательства в связь и управление), что могло бы нарушить работу систем и исказить процесс принятия военных решений во время кризиса и в других оперативных условиях.

Во время конфликта эти технологии и темпы операций, которые они обеспечивают, вероятно, значительно сократят время принятия решений, ограничат возможности человеческого понимания и потребуют машинной скорости

реагирования. Поскольку многие средства ИИ функционируют по принципу «черного ящика», человек часто не в состоянии понять, как и почему был получен тот или иной результат. Операции, осуществляемые при поддержке ИИ, могут стать причиной непредсказуемого и непрозрачного поведения, что затруднит получение точных выводов и суждений о намерениях противника, может привести к их неверному толкованию или спровоцировать непреднамеренные последствия. Может возникнуть ситуация, когда операторы будут чрезмерно доверять алгоритмическим результатам, не имея полного представления об исходных предпосылках, ограничениях и недостатках систем ИИ. В отсутствие соответствующих гарантий, норм и протоколов системы, функционирующие при поддержке ИИ, могли бы усугубить риск недоразумений, просчетов и непреднамеренной эскалации.

Широкое распространение передовых средств или инструментов ИИ и других технологий двойного назначения, вероятно, повышает риски распространения и разработки новых вооружений оружия государственными и негосударственными субъектами. Кроме того, ИИ может использоваться для дополнения или усиления попыток дезинформации, призванных возбудить враждебное отношение к странам, что может вызвать конфликт и эскалацию напряженности.

#### **Приверженность Соединенного Королевства безопасному и ответственному применению искусственного интеллекта в военной области**

Соединенное Королевство признает, что ИИ порождает глубокие опасения по поводу справедливости, предвзятости, надежности и характера ответственности и подотчетности человека, особенно в военном контексте. Хотя государства уже давно используют новые технологии и будут и далее опираться на давно установленные правовые режимы, режимы безопасности и регулирования, мы должны признать, что характер ИИ порождает особые проблемы, и признать важность решительной демонстрации нашего ответственного и надежного подхода.

Соединенное Королевство сформулировало свое обязательство в отношении безопасного и ответственного применения ИИ в своей Оборонной стратегии по вопросам ИИ и связанных с ней этических принципах применения ИИ. Эти этические принципы применения ИИ, изложенные в политике Соединенного Королевства под названием «Амбициозный, безопасный, ответственный», устанавливают этические рамки, включающие человекоцентричность, ответственность, понимание, предвзятость, уменьшение вреда и надежность. В Совместном издании вооруженных сил под названием «Надежный искусственный интеллект (ИИ) в сфере обороны», которое вышло в ноябре 2024 года, содержатся четкие указания для служащих Министерства обороны и других ведомств относительно реализации этих этических принципов применения ИИ для обеспечения надежных, прочных и эффективных услуг и возможностей с использованием ИИ.

Посредством своих этических принципов применения ИИ Соединенное Королевство стремится укрепить доверие к технологиям ИИ и их применению, реализовать весь потенциал взаимодействия человека и машины, уменьшая при этом риски, связанные с его применением, ненадлежащим применением или отказом от применения, и предотвращая непреднамеренные последствия. Такой подход позволяет Соединенному Королевству использовать инновации и творческий потенциал в рамках оборонной и промышленной отраслей для обеспечения широкого внедрения решений с поддержкой ИИ.

Правительство Соединенного Королевства четко указало, что любое использование Соединенным Королевством ИИ для совершенствования процессов в сфере обороны, улучшения оборонных систем или наращивания военного потенциала регулируется национальным законодательством и нормами международного права. Вооруженные силы Соединенного Королевства неизменно стремятся соблюдать свои юридические обязательства по всему спектру деятельности, начиная с трудового законодательства и заканчивая вопросами конфиденциальности, закупок и права вооруженных конфликтов, также известного как международное гуманитарное право. В стране приняты надежные методы и процессы, обеспечивающие соблюдение закона в рамках осуществляемых мероприятий и физическими лицами. Эти методы и процессы уже применяются и будут применяться к средствам с поддержкой ИИ. Развертывание средств с поддержкой ИИ в вооруженном конфликте должно полностью соответствовать нормам международного гуманитарного права и удовлетворять его четырем основным принципам: избирательность, необходимость, гуманность и соразмерность. Мы четко понимаем, что применение любой системы или вооружений, не отвечающих этим основополагающим принципам, станет нарушением норм международного права.

Кроме того, критически важное значение имеют ответственность и подотчетность человека, обеспечиваемые участием операторов с учетом конкретных условий. Такое участие операторов с учетом конкретных условий необходимо для того, чтобы удовлетворять требованиям наших стратегий и обеспечивать соблюдение этических принципов и обязательств по международному гуманитарному праву. Характер участия человека будет варьироваться в зависимости от характера задействуемых средств, оперативной обстановки и условий применения. Соединенное Королевство будет обеспечивать постоянный контроль человека за проводимой страной политикой в отношении ядерного оружия.

### **Вклад Соединенного Королевства в международные инициативы**

Глобальная стабильность требует энергичного, но ответственного развития военного ИИ. Понимание международным сообществом рисков, гарантий и стандартов, связанных с применением ИИ в военном контексте, продолжает развиваться. Поскольку такие риски по своей природе носят международный характер, они требуют принятия глобальных мер реагирования.

Соединенное Королевство находится в авангарде международных усилий по поддержке безопасного и ответственного развития и применения ИИ. Соединенное Королевство гордится тем, что стало принимающей страной первого саммита по вопросам безопасности ИИ, на котором была принята Блетчлийская декларация о безопасности ИИ, и сыграло свою роль в подготовке Международного доклада по вопросу безопасности ИИ, который был опубликован в феврале 2025 года и стал первым в мире всеобъемлющим обобщением современной литературы о рисках и возможностях передовых систем ИИ, способствовавшим углублению понимания аспектов, имеющих решающее значение для информационного обеспечения международных дискуссий по таким вопросам, как использование ИИ в интересах мира и безопасности. Мы поддерживаем усилия в рамках Глобального цифрового договора, направленные на устранение цифрового разрыва и укрепление международного регулирования в вопросах использования ИИ на благо человечества.

Соединенное Королевство активно поддерживает международные инициативы, направленные на стимулирование действий в связи с применением в военной области. Мы поддерживаем работу таких организаций, как «РЕНД Европа», Калифорнийский университет в Беркли и Глобальная комиссия по

ответственному применению искусственного интеллекта в военной области, преследуя цель собрать воедино различных широко известных экспертов для изучения этих вопросов, осмысления последних достижений, определения путей продвижения вперед для политиков и выработки для них практически осуществимых рекомендаций.

Соединенное Королевство продолжает активно участвовать в международном диалоге по вопросам обороны и безопасности, связанным с ИИ, и делится своим опытом разработки и осуществления безопасных и ответственных подходов к внедрению ИИ в военной области. Соединенное Королевство приветствует прогресс, достигнутый благодаря таким инициативам, как проведение саммитов по ответственному использованию искусственного интеллекта в военной сфере, одним из организаторов которого Соединенное Королевство выступило в 2024 году, и принятие под руководством Соединенных Штатов Политической декларации об ответственном применении искусственного интеллекта и автономных систем в военных целях, чтобы лучше понять возможности и стратегические риски, а также способы их устранения путем принятия соответствующих мер в поддержку безопасного и ответственного применения ИИ. Этика и гарантии в связи с применением ИИ представляют собой динамично развивающиеся сферы, которые требуют постоянного участия и сотрудничества и применения итерационного подхода.

### **Перспективы**

Соединенное Королевство рассчитывает закрепить достигнутый на сегодняшний день прогресс в рамках существующих процессов, в том числе по линии обсуждений в Организации Объединенных Наций, основанных на докладе Генерального секретаря и нацеленных на принятие конкретных мер. С учетом характера применения ИИ в военном контексте будет крайне важно применять инклюзивный подход с участием многих заинтересованных сторон с опорой на технические, военные и правовые знания и опыт государственных ведомств, отраслевых и научных кругов и гражданского общества.

Несмотря на обилие информации, наше коллективное понимание военных применений и последствий по-прежнему находится на низком уровне и сохраняются существенные пробелы в знаниях и понимании природы и возможностей ИИ. Необходимо продолжать работу по наращиванию потенциала государств, углублению нашего коллективного понимания последствий и потенциальных рисков и проблем применения военного ИИ на стратегическом уровне и выработке общеприемлемой терминологии, чтобы можно было проводить конструктивные дискуссии. В ходе обсуждений необходимо сделать особый упор на поиске конкретных, эффективных и надлежащих мер и методов, которые могли бы способствовать устранению рисков, включая гарантии и нормы поведения, новые каналы связи и механизмы обеспечения транспарентности ради уменьшения риска недоразумений, обновленные доктрины, меры укрепления доверия и соглашения о контроле над вооружениями с учетом последствий военного применения ИИ.



## В. Европейский союз

[Подлинный текст на английском языке]  
[11 апреля 2025 года]

Европейский союз приветствует эту возможность изложить свои взгляды на проблемы и возможности в сфере обеспечения международного мира и безопасности, создаваемых применением искусственного интеллекта (ИИ) в военной области, в соответствии с резолюцией 79/239 Генеральной Ассамблеи от 24 декабря 2024 года.

Прежде всего, Европейский союз хотел бы напомнить о своей давно сформулированной позиции, согласно которой применение ИИ в военной области должно соответствовать нормам международного права, в особенности Уставу Организации Объединенных Наций, нормам международного гуманитарного права и международного права прав человека.

Европейский союз хотел бы напомнить и о другой давней позиции, а именно о том, что человеческое суждение и контроль над применением силы должны сохраняться во всех случаях. Кроме того, человек должен сохранять ответственность и подотчетность в вопросах применения ИИ в военной области, чтобы обеспечивать ответственное применение такой технологии.

Европейский союз признает, что применение ИИ в военных системах порождает как возможности, так и проблемы. Развитие ИИ происходит настолько быстро, что на данный момент невозможно предсказать все преимущества и риски.

В этой связи Европейский союз приветствует постоянное внимание к этому вопросу со стороны Организации Объединенных Наций, а также обсуждения на соответствующих международных форумах. В этой связи Европейский союз особенно высоко оценивает продолжение процесса по тематике ответственного искусственного интеллекта в военной области, начатого в Нидерландах в 2023 году с первого саммита по ответственному использованию искусственного интеллекта в военной сфере, за которым последовал саммит по ответственному использованию искусственного интеллекта в военной сфере, прошедший в Республике Корея в 2024 году. Европейский союз приветствует продолжение этого процесса в рамках саммита по ответственному использованию искусственного интеллекта в военной сфере, который состоится в 2025 году в Испании, и выражает благодарность Испании за организацию этого следующего саммита.

Европейский союз отмечает, что Призыв к действиям по ответственному применению искусственного интеллекта в военной области 2023 года, а также План действий по ответственному использованию искусственного интеллекта в военной области 2024 года были одобрены всеми государствами — членами Европейского союза. Европейский союз считает, что концепция саммита по ответственному использованию искусственного интеллекта в военной сфере, предусматривающая многосторонние и инклюзивные процессы в рассмотрении вопроса об ответственном применении ИИ в военных целях, представляет собой многообещающий подход. В этой связи Европейский союз признает ценность других недавних мероприятий, таких как Международный саммит по ответственному интеллекту и Саммит действий по ответственному интеллекту, который был проведен во Франции 10 и 11 февраля 2025 года. Европейский союз также отмечает работу в рамках Политической декларации об ответственном применении искусственного интеллекта и автономных систем в военных целях в качестве ценного вклада в более широкую международную дискуссию о последствиях применения ИИ в интересах обеспечения международного мира и безопасности.

Европейский союз считает, что итоговые документы об ответственном применении искусственного интеллекта в военной области и Политическая декларация, которую подписали все государства — члены Европейского союза, дополняют друг друга и имеют очень большое значение для дальнейшего развития глобального мышления, управления и практических решений в сфере ответственного применения ИИ в военной области.

Европейский союз признает, что применение ИИ в военной области имеет военные преимущества. Они заключаются, в частности, в оперативности, масштабах и точности военных действий. ИИ может обеспечивать тактическое преимущество путем организации и предварительной обработки огромных массивов данных, поступающих от систем наблюдения и вооружений, беспилотных аппаратов и спутниковых снимков, что позволит операторам принимать решения быстрее и качественнее. Приложения ИИ могут сократить расходы благодаря совершенствованию материально-технического снабжения или обслуживания материальной части посредством управления прогнозируемым техническим обслуживанием. Кроме того, ИИ может обеспечивать возможность увеличения дистанции и повышения точности при проведении военных операций в условиях неопределенности.

В то же время те же самые преимущества с точки зрения оперативности и масштаба, которые обеспечивает применение ИИ в военной области, создают и проблемы. ИИ ускоряет цикл наблюдения, оценки обстановки, принятия решения и нанесения удара. Нарастание скоростных возможностей и расширение масштаба могут стать причиной заблуждений в результате несоответствия между военными намерениями и анализом, проводимым системами под управлением ИИ. Поэтому ИИ может непреднамеренно способствовать эскалации. Оперативность также создает проблемы с точки зрения необходимости сохранения участия человека в принятии решений и контроля за применением силы.

В этой связи Европейский союз подчеркивает важность международного сотрудничества, направленного на изучение последствий применения ИИ в военной области и возможных механизмов регулирования.

## Annex II

### **Replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry<sup>1</sup>**

#### **A. International and regional organizations**

##### **African Commission on Human and Peoples' Rights**

[11 April 2025]

##### **I. Introduction**

The African Commission on Human and Peoples' Rights (the African Commission), as the premier treaty-based human and peoples' rights body of the African Union (AU), is entrusted with the mandate of promoting and protecting human and peoples' rights in Africa under the African Charter on Human and Peoples' Rights (African Charter). In the African Commission's study on Addressing Human Rights Issues in Conflict Situations, the African Commission's Focal Point who led the study observed that 'it is ... in conflict and crisis situations that the most egregious violations and abuses of rights are perpetrated...With the changes in the nature of conflicts and the attendant heightened threat to human and peoples' rights, there is a greater need for the human rights system to pay increasing attention to and provide effective responses to the challenges that these new dynamics present to the protection and observance of rights.' In the current context, one of the major new dynamics that carries serious implications for peace and security and therefore human and peoples' rights relate to Artificial Intelligence (AI) and in particular its rapid development and use in the military domain.

During its 1214th meeting, the AU Peace and Security Council (PSC), in requesting the AU Commission to conduct a study to assess the adverse impact of AI on peace and security, underscored the necessity of ensuring African perspectives in shaping global AI governance frameworks. Against this background and having regard to its work on AI and other technologies and human and peoples' rights<sup>2</sup> and human rights in peace and security, the African Commission is pleased to share its views in response to the invitation of the Secretary-General for submission of inputs on AI in the military domain and its implications for international peace and security.<sup>3</sup>

##### **II. AI in the military domain and peace and security**

The development and use of AI technologies in the military domain particularly to automate military functions such as surveillance, targeting, and the deployment of

<sup>1</sup> In accordance with operative paragraph 8 of General Assembly resolution 79/239, the replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry are included in the original language received. The Secretary-General remains committed to multilingualism as a core value of the United Nations.

<sup>2</sup> Resolution ACHPR/Res. 473 (EXT.OS/ XXXI) 2021 on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa, available at <https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-humanand-peoples-rights-and-art>.

<sup>3</sup> The Focal Point of the African Commission on its study on human and peoples' rights and AI, robotics and other technologies acknowledges with appreciation the contribution of Professor Thompson Chengeta, who is the consultant providing technical assistance in the development of the study, through the Centre for Human Rights, University of Pretoria.

lethal force have far reaching consequences for peace and security and hence for human and peoples' rights. The AU Continental AI Strategy, endorsed during the 44th Extraordinary Session of the Executive Council of the African Union, highlights AI governance and regulatory challenges, particularly in military applications, warning that AI could exacerbate conflicts through inaccurate predictions or deployment of autonomous weapon systems. Additionally, the framework raises concern about disinformation, misinformation, cybersecurity threats, and military risks.

From the perspective of the development and use of AI in the military domain, peace and security should not be seen just from the perspective only of what it means for stability of states and societies. Beyond its conception under the UN Charter and public international law associated with friendly relations of states, peace and security is also a fundamental right of all peoples. The African Charter thus stipulates that 'All peoples shall have the right to national and international peace and security. The principles of solidarity and friendly relations implicitly affirmed by the Charter of the United Nations and reaffirmed by that of the Organization of African Unity shall govern relations between States.'<sup>4</sup>

The framing of peace and security as a right of peoples compels states to assess and govern the development and deployment of AI technologies in the military domain through a human rights lens that prioritises the prevention of harm, suffering, and injustice. Together international law conception of peace and security, it places an affirmative duty on states to ensure that AI systems do not contribute to conflict, perpetuate structural inequalities, or violate the rights and dignity of individuals and communities. By embedding peace and security within the framework of human rights, states are not only accountable for avoiding direct acts of aggression, but also for proactively creating and maintaining environments in which human flourishing, security, and justice are protected from the potentially disruptive or harmful impacts of emerging military technologies.

The implication of AI in the military domain to peace and security, farmed comprehensively, thus goes beyond how it shapes the obligation of states for non-aggression. It also covers how algorithm-driven systems may dehumanise individuals, introduce bias, and lead to unaccountable or disproportionate harm. It raises critical questions about the erosion of human oversight, the potential for unlawful killings or violations of international humanitarian law, and the targeting of vulnerable or marginalised populations.

By transforming military capabilities, the application of AI in the military domain can also have implications for peace and security by heightening tendencies for engaging in hostilities. The resultant escalation of tension and violence will be inimical not only to stability and peace between and within states but also most importantly carries more adverse consequences for the development needs of the less developed parts of the world such as Africa. While AI may contribute to advancing the development needs of Africa, its development and use in the military domain can have devastating consequences for development detrimental in particular to the right to development enshrined in Article 22 of the African Charter.<sup>5</sup>

This link between peace and development is also central to the Sustainable Development Goals (SDGs), especially SDG 16, which promotes peace, justice, and strong institutions. Without peace and security, sustainable development cannot be

<sup>4</sup> Article 23(1) of the African Charter.

<sup>5</sup> All peoples shall have the right to their economic, social and cultural development with due regard to their freedom and identity and in the equal enjoyment of the common heritage of mankind.

achieved. Recognising this link is critical in the governance of military AI, as the militarisation of AI can aggravate instability, particularly in fragile regions, and undermine Africa's developmental aspirations. By reaffirming the interconnectedness of peace and development, the African Commission calls for a governance approach that upholds peace as both a human right and a developmental imperative.

### **III. The need for a human and peoples' rights-based regulation of the development and use of AI in the military domain**

Given the ways in which the use of AI in the military domain transforms the conduct of hostilities and how the development of AI relies on the extraction of natural resources particularly critical minerals such as rare earth minerals, it is the submission of the African Commission that both the process of extraction of resources in the development of AI in the military domain and the use of AI in the military domain need to be in full compliance with human and peoples' rights standards and international law principles, including international humanitarian law.

First and foremost, it is of paramount significance that the development and use of AI in the military domain complies with the right to peace and security enshrined in Article 23 of the African Charter on Human and Peoples' Rights. As a right that is born out of the recognition of the inseparability of the enjoyment of other human rights states from peace and security, this right entails that the use of AI in the military domain should be consistent with the international law prohibition of the use of force enshrined in the UN Charter and the Constitutive Act of the African Union.

Second, the use of AI technologies in conflict settings need to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy. These requirements apply irrespective of whether the context in which the use of AI in the military domain relates to international armed conflicts or non-international armed conflicts. As established in the African Commission's study,<sup>6</sup> parties to conflict are obliged to observe human rights standards where such conflicts do not meet the IHL threshold of armed conflict. As such, those who use AI technologies in conflict situations that do not meet the IHL threshold of armed conflict are legally obliged to respect and ensure respect for the human and peoples' rights standards established under treaty and customary international human rights law.

Third, the development of AI in the military domain and the use AI technologies in hostilities need to comply with the principle of transparency. This is fundamental because it is the basis for ensuring effective regulation of the development and use of AI in the military domain and for compliance with applicable human rights and international law standards. Additionally, transparency is critical for ensuring compliance with the obligation for respecting the dignity, privacy and data protection of individuals. The principle of transparency is also a pre-requisite for addressing some of the concerns that arise from use of AI in the military domain including bias (owing to the source and type of data used) and explainability. Transparency is also critical not only with the development of AI in the military domain but also with respect to the transfer of AI technologies in the military domain.

Fourth, from the perspective of human and peoples' rights and IHL, the other standard key to human rights and international law-based regulation of the development and use of AI concerns accountability. In the event of the occurrence of violations of human and peoples' rights standards or IHL principles from the development and use of AI in the military domain, there has to be both institutional

<sup>6</sup> ACHPR, Addressing human rights issues in conflict situations, <https://achpr.au.int/en/node/895>.

and individual accountability. Accountability in this instance encompasses not only the measures that are taken against perpetrators but also the remedial steps that need to be put in place for redressing victims.

Firth, building and sharing of technical knowhow critical to ensuring regulation by states is the other principle. Recent developments including the jamming of GPS systems affecting flights reported in Eastern DRC and the deployment by the Islamic State of West Africa of armed drones, highlight not only the need for effective regulation but also the need for developing the requisite infrastructure and technical capacity for ensuring effective regulation.

#### **IV. The link between the development of AI in the military domain and Africa's natural resources and its implications for peace and security**

The African Commission is also of the view that when discussing peace and security, stakeholders must be aware of the link between development of military AI, Africa's natural resources – particularly critical minerals – and the notion of peace and security. Article 21(1) of the African Charter on Human and Peoples' Rights affirms: "All peoples shall freely dispose of their wealth and natural resources. This right shall be exercised in the exclusive interest of the people. In no case shall a people be deprived of it."<sup>7</sup> Article 21(5) further provides that "States parties to the present Charter shall undertake to eliminate all forms of foreign economic exploitation particularly that practised by international monopolies so as to enable their peoples to fully benefit from the advantages derived from their national resources."<sup>8</sup>

This provision is particularly important in the context of military AI, which depends heavily on critical minerals such as cobalt, lithium, and rare earth elements – resources abundantly found in Africa. The 2024 Report of the Chairperson of the African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations, stressed the "significance of critical minerals for new and emerging technologies" and highlighted that Africa has been burdened by a "resource curse phenomenon."<sup>9</sup> The report of the Chairperson noted that "extraction of minerals and other resources not only fuels but also at times becomes the site where contestation over whose control and use triggers conflicts. In some instances, this has created a vicious cycle of insecurity and violence, a condition that not only leads to major human and peoples' rights violations but also the perpetuation of a vacuum of effective governance and the concomitant exploitative, socially and environmentally costly extraction of the resources of the continent."<sup>10</sup>

Therefore, governance of military AI must not only ensure the legal use of force but also address the exploitative chains of extraction that power such technologies. This requires strict oversight, equitable benefit sharing, and regional solidarity to prevent Africa's resources from being used to fuel further conflict and inequality.

#### **V. Conclusion**

The African Commission is of the view that the development and use of AI in the military domain carries far reaching consequences for international peace and security in general and for less developed parts of the world such as in Africa that historically suffered violations and remain vulnerable to the adverse impacts of the development and use of AI in the military domain without robust and effective legal

<sup>7</sup> Article 21(1) of the African Charter.

<sup>8</sup> Article 21(5) of the African Charter.

<sup>9</sup> African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations (2024), <https://achpr.au.int/en/intersession-activity-reports/extractive-industries-environment-and-human-rights-violations> (accessed 08 April 2025).

<sup>10</sup> As above.

regime for such development and use in the military domain. The African Commission affirms that the development and use of AI in the military domain needs to be regulated on the basis of international law, human and peoples' rights and international humanitarian law standards with particular regard to the development and peace and security interests and human and peoples' rights needs of less developed parts of the world.

More specifically, beyond and above the right to peace and security, the governance of AI in the military domain needs to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy, the principles of transparency, accountability and redress for victims and the obligation to build and share technical knowhow necessary for enabling societies to avert the risks that the development and use of AI in the military domain carries for peace and security. Only by ensuring that the development and use of military AI are aligned with international legal standards including those relating to the right to peace and security, the right to development, the right to privacy and protection of personal data, the right to remedy and the responsibility for exercising human control, the right to and control over natural resources and by addressing the structural inequities underpinning global technological advancement, can states uphold their duties to their peoples and advance genuine peace, justice, and security in relation to the development and use of AI in the military domain.

## B. International Committee of the Red Cross

[19 March 2024]

### Summary

The full submission is available at: <https://www.icrc.org/en/article/artificial-intelligence-military-domain-icrc-submits-recommendations-un-secretary-general>.

The International Committee of the Red Cross (ICRC) welcomes the opportunity to submit its views for consideration by the United Nations Secretary-General, in accordance with resolution 79/239.

The recommendations that the ICRC makes in this submission are in line with its long-standing mandate and practice of promoting respect for and the development of IHL, including its application to new technologies of warfare. This submission is intended to support States in ensuring that military applications of AI comply with existing legal frameworks and, where necessary, identifying areas where additional legal, policy, or operational measures may be required.

### 1. Normative proposals: Reaffirming existing IHL as the starting point

The ICRC has consistently emphasized that, while IHL does not explicitly prohibit or regulate the use of AI in military applications, it does restrict its development and use, and places strict constraints on AI when it is integrated into weapon systems or used in some way to conduct warfare.<sup>1</sup>

Existing and emerging normative proposals on the military application of AI should build upon established international legal frameworks and mechanisms, including IHL. Where necessary, these frameworks can be reinforced through the development of additional legal instruments, operational guidance or policy measures

<sup>1</sup> This has also been affirmed by States, including in the UN General Assembly with Resolution 79/239.

to address specific risks or challenges posed by emerging technologies. The form and content of such measures may vary depending on the specific use case. The ICRC encourages the international community to engage in concrete discussions on particular applications of AI in the military domain and to prioritize consideration of those that pose the greatest risks to people affected by armed conflicts.

## 2. A Human-centred Approach to military AI

In line with the resolution, the ICRC advocates for a human-centred approach to the development and use of AI in armed conflict.<sup>2</sup> This approach has at least two key dimensions: first, ensuring a focus on the humans who may be affected by the use of AI; and second, emphasizing the obligations and responsibilities of the humans using or ordering the use of AI in military operations.

Despite the growing development of AI-related technologies in the military domain, IHL requires individuals to make legal determinations. Humans must, for instance, determine the lawfulness of attacks that they plan, decide upon or execute, and they remain accountable for those determinations. The ICRC considers that human judgement is crucial for reducing humanitarian risks, addressing ethical concerns and ensuring compliance with IHL. Accordingly, while certain technical tasks may be carried out by machine processes, it is not the system itself that must comply with the law, but the humans using it.<sup>3</sup>

This does not mean that commanders and combatants cannot or should not use tools, including AI-decision-support systems. However, these tools must only be designed and used to support, rather than hinder or replace, human decision-making.<sup>4</sup> Further, States and parties to armed conflicts must ensure that human control and judgement are preserved in decisions that pose risks to the life and dignity of people affected by armed conflict. This is essential for ensuring respect for applicable laws, including IHL, and upholding ethical standards.<sup>5</sup>

## 3. Specific Applications of ai in the military domain

The ICRC has identified three specific applications of AI in the military domain that pose particularly significant risks to those affected by armed conflict:

### 1. *AI in Autonomous Weapon Systems*

Resolution 79/239 acknowledges the increasing integration of AI into weapons and weapon systems, a development that raises significant legal and humanitarian concerns. The integration of AI, particularly machine learning (ML) techniques, into autonomous weapon systems (AWS) exacerbates existing challenges posed by AWS in ensuring compliance with IHL. In particular, it increases difficulties for human users to understand, predict, and control the system's functioning and effects.

Users of AWS must be able to, with a reasonable degree of certainty, predict the effects of that weapon in order to determine whether it can be directed at a specific military objective, and take steps to limit those predicted effects, as required by IHL.

<sup>2</sup> ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021).

<sup>3</sup> ICRC, International Humanitarian Law and the Challenges of Contemporary Armed Conflicts: Building a Culture of Compliance for IHL to Protect Humanity in Today's and Future Conflicts (IHL Challenges Report), 2024, p. 61.

<sup>4</sup> *Ibid.*; ICRC, IHL Challenges Report – Chapter 2: Contemporary and future challenges in the conduct of hostilities, 2019, p. 32.

<sup>5</sup> ICRC, Decisions, Decisions, Decisions: computation and Artificial Intelligence in military decision-making, ICRC, 2024, p. 8.



This entails the ability to understand the functioning of the AWS: the nature and functioning of its sensors, the definition of its target profile and the potential effects in the circumstances of use, including any risk of error or malfunction. This is particularly relevant for AWS that function in opaque ways (the “black box” challenge), such as AWS relying on AI techniques, which prevent the human user from being able to understand, predict or explain the system’s output. This impossibility effectively results in a lack of control over the weapon’s effects, rendering it indiscriminate by nature.

In this regard, we reiterate the joint call made by the ICRC President, with the UN secretary-general,<sup>6</sup> for new, legally binding rules prohibiting certain AWS and constraining the use of others.<sup>7</sup> In particular, we recommend a prohibition on

- unpredictable autonomous weapons – those that, due to their design or the circumstances and manner of use, do not allow a human user to understand, explain or predict the system’s functioning and effects;
- autonomous weapons designed or used to target humans directly. This is required because of the significant risk of IHL violations and the unacceptability of anti-personnel autonomous weapons from an ethical perspective.<sup>8</sup>

The ICRC supports all efforts by States to urgently adopt a legally binding instrument to regulate AWS, in whichever forum they choose.<sup>9</sup> The integration of AI into AWS should also be considered when discussing normative proposals on military applications of AI. Doing so is essential to ensure a consistent and comprehensive approach to the regulation of military AI, to avoid normative gaps, and to effectively address the serious legal, ethical, and humanitarian risks that are exacerbated by the integration of AI into AWS. In this regard, the ICRC considers it important that binding prohibitions and restrictions on AWS, including AWS that incorporate AI, are integrated into broader discussions on the governance of military AI.

## 2. *AI in Military Decision-Making*

AI decision-support systems (AI-DSS) are computerised tools that bring together data sources – such as satellite imagery, sensor data, social media feeds or mobile phone signals – and draw on them to present analyses, recommendations and predictions to decision makers.

The use of AI-DSS raise concerns related to system functioning, data quality, and human-machine interaction. These systems risk increasing the rate of unforeseen errors, perpetuating problematic biases – particularly those based on age, gender, ethnicity, or disability, and making it difficult for the users to understand how and why the system generates its output from a given input.

Generally, AI-based systems will perform better when given well-defined goals and access to representative and high-quality data. However, armed conflict environments are marked by uncertainty, volatility, and deliberate deception techniques by adversaries, which makes it extremely difficult to obtain reliable or transferable data. Even where good data exists, it may not reflect the specific operational or humanitarian dynamics of a particular context.<sup>10</sup> Moreover, for AI

<sup>6</sup> ICRC, Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 2023.

<sup>7</sup> ICRC, ICRC Submission on AWS to the UN Secretary-General, 2024, p. 6.

<sup>8</sup> *Ibid.*

<sup>9</sup> *Ibid.*

<sup>10</sup> ICRC, IHL Challenges Report, 2024, pp. 64–65; ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021); ICRC, Decisions, Decisions,

systems that rely on training data, the utility of those data can rapidly diminish once a conflict begins. Parties to armed conflicts will continuously seek to maintain the initiative and operate in a manner that is not anticipated by their adversary, adapting their strategies and tactics accordingly. This can fundamentally alter the environment in which the system was expected to operate, making the original data no longer representative of the new operational conditions. In such cases, the system's outputs may become unreliable, and the AI model may require re-evaluation or retraining in order to remain fit for purpose.

Human interaction with these systems raises further concerns, such as “automation bias” – a propensity to rely on machine outputs even when other available information may call those outputs into question – which is particularly pronounced in high-pressure or stressful environments like in armed conflicts.<sup>11</sup> Taken together, these factors can hamper a user's ability to scrutinize the information available. The practical consequence might be, for instance, that someone plans, decides upon or launches an attack based solely on an AI-DSS's output, thereby effectively serving as a human rubber stamp rather than assessing the lawfulness of the attack by considering all the information reasonably available including the AI-DSS output.<sup>12</sup>

On the positive side, the careful use of AI-based systems may facilitate quicker and more comprehensive information analysis, which can support decisions in a way that enhances IHL compliance and minimizes risks for civilians. In the context of urban warfare in particular, the ICRC has recommended that online open-source repositories should be used to gather information about the presence of civilians and civilian objects.<sup>13</sup> Importantly, IHL imposes obligations to take constant care to spare the civilian population and to take all feasible precautions in attack. Therefore, in developing and using AI-DSS, armed forces should be considering not only how such tools can assist them to achieve military objectives with less civilian harm, but also how they might be designed and used specifically to protect civilians. However, the important point is that these computer outputs can inform but must not displace the need for legal determinations.

Beyond targeting decisions, militaries are also exploring the use of AI to support other operations traditionally carried out by humans, including detention operations. While technology deployed responsibly and with robust human oversight can contribute to IHL compliance, it also carries risks including bias, lack of transparency, and faulty programming and analysis, all of which can undermine compliance with IHL.<sup>14</sup>

To support efforts by States and other actors to ensure that military uses of AI-DSS remain consistent with IHL and humanitarian principles, the ICRC has formulated a non-exhaustive set of preliminary recommendations relating to the development and use of AI-DSS in armed conflict. They focus on 1) ensuring human control and judgement; 2) system design requirements; 3) testing, evaluation, verification and validation; 4) legal reviews; 5) operational constraints on use; 6) user

---

Decisions: Computation and Artificial Intelligence in Military Decision-Making, ICRC, 2024, pp. 31 and 54.

<sup>11</sup> ICRC and the Geneva Academy, Artificial Intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts: Current Developments and Potential Implications, ICRC, 2024, p. 17.

<sup>12</sup> ICRC, IHL Challenges Report, 2024, p. 65.

<sup>13</sup> *Ibid.*, p. 66; ICRC, Reducing Civilian Harm in Urban Warfare: A Handbook for Armed Groups, 2023, p. 15.

<sup>14</sup> ICRC, IHL Challenges Report, 2024, p. 22.

training; 7) after-action reviews; and 8) accountability, among others. The recommendations are annexed to the full version of this submission.

### 3. *AI in Information and Communications Technologies*

AI is expected to change how actors defend against and conduct information and communications technology (ICT) activities, including in armed conflict. In particular, States have noted with concern that the use of AI and other emerging technologies in malicious ICT activities may further increase their scale and speed, as well as the harm they may cause.<sup>15</sup> For example, AI enables tools to identify and develop exploits for new vulnerabilities in software or networks, or to conduct harmful ICT activities autonomously, whether in offence or in defence. The ICRC is concerned that this could increase the risks of indiscriminate attacks, incidental civilian harm, including damage to critical civilian infrastructure, as well as the uncontrolled escalation of conflict, particularly in complex and interconnected digital environments.<sup>16</sup>

Similarly, information or psychological operations are not a new feature of armed conflicts; however, AI is changing how information is created and spread. AI-enabled systems, particularly generative AI, have been widely used to produce harmful content – text, audio, photos and video – which is increasingly difficult to distinguish from authentic, original content.<sup>17</sup> The ICRC is concerned about the consequences for civilians that might result from the creation and spread of such information through ICT, including information that contributes to or encourages violence, causes lasting psychological harm, undermines access to essential services or disrupts the operations of humanitarian organizations.

In light of these concerns, the ICRC underlines the importance of applying existing international law, including IHL, to the use of AI in ICT activities. The ICRC urges States to ensure that the development and use of AI-supported ICT activities respect the protections afforded to civilians and civilian infrastructure in armed conflict. Moreover, in light of the emergence of increasingly autonomous ICT capabilities, the ICRC further encourages States to address the serious challenges posed by these tools, particularly by considering whether existing international law, including IHL, provides sufficient safeguards against the harm such tools can cause, or whether additional limits are needed.

### 4. **Conclusion**

The ICRC is grateful for the opportunity to share its above views and recommendations on ways to address the challenges and concerns raised by AI for the secretary-general's consideration, and stands ready to contribute further to assist States in taking effective action to address the risks posed by AI applications in the military domain.

<sup>15</sup> 34th International Conference of the Red Cross and Red Crescent, Resolution 2 “Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict”, 2024.

<sup>16</sup> ICRC, IHL Challenges Report, 2024, pp. 66–67.

<sup>17</sup> *Ibid.*, pp. 58–59.

## C. Civil society

### Autonorms

[10 April 2025]

The following is the AutoNorms project's submission pursuant to Resolution [79/239](#) on "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the United Nations General Assembly on 24 December 2024. The resolution requests the UN Secretary-General to seek views, including those of Member States, civil society, the scientific community and industry, on "opportunities and challenges posed by the application of artificial intelligence in the military domain, **with specific focus on areas other than lethal autonomous weapons systems**". The AutoNorms team welcomes the opportunity for representatives of academia to submit their views on this important and timely topic.

The AutoNorms project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852123). Led by Professor Ingvild Bode and hosted by the Center for War Studies at the University of Southern Denmark, the project examines how the integration of artificial intelligence (AI) technologies into weapon systems and military targeting shapes international norms governing the use of force.<sup>1</sup>

### Introduction

Over the past 2-3 years, the international debate about applications of AI in the military domain has been characterized by two significant, near-simultaneous changes. First, there has been a **move away from its predominant focus on autonomous or AI technologies in weapon systems** towards considering AI technologies across a wider range of military decision-making tasks, especially in relation to targeting. To reflect this move, this submission focuses on the employment of **AI-based decision support systems (AI DSS)**, or systems that are meant to be used as tools to directly or indirectly inform the complex process of use-of-force decision-making, for example, by analyzing large volumes of data, recognizing patterns within the data, predicting scenarios, or recommending potential courses of action to human decision makers.

Second, there has been a **growing emphasis on human-machine interaction** in the context of using AI in the military domain.<sup>2</sup> This emphasis results from the broad recognition that, even when humans are 'in' or 'on' the loop of targeting decision-making, they need to exercise a sufficient level of oversight, control, and agency over the targeting process. Human oversight is a governance principle featuring prominently across various international initiatives, including [A/RES/79/239](#). However, **dynamics of human-machine interaction as part of the use of AI DSS both introduce new issues and solidify existing sets of challenges that require governance attention**. Our submission highlights these challenges and the need to ensure the exercise of human oversight and agency throughout the full targeting decision-making spectrum. It is structured in three parts, starting with

<sup>1</sup> The members of the AutoNorms team are Professor Ingvild Bode, Dr Hendrik Huelss, Dr Anna Nadibaidze, Dr Guangyu Qiao-Franco, and Dr Qiaochu Zhang. The AutoNorms project is based at the Center for War Studies, University of Southern Denmark, Odense, Denmark. For more information, please visit our website: [www.autonorms.eu](http://www.autonorms.eu).

<sup>2</sup> Ingvild Bode and Anna Nadibaidze, "Symposium on Military AI and the Law of Armed Conflict: Human-Machine Interaction in the Military Domain and the Responsible AI Framework," *Opinio Juris*, April 4, 2024, <https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-human-machine-interaction-in-the-military-domain-and-the-responsible-ai-framework/>.

explicating challenges of human-machine interaction, then commenting on the relative under-development of the international debate about AI DSS, and finally, sketching a way forward.

### Challenges of human-machine interaction in the use of AI DSS

The use of AI DSS involves various dynamics of human-machine interaction because military personnel such as operators and intelligence analysts routinely and increasingly interact with a network of AI systems throughout the targeting process. These interactions involve multiple challenges **which have the potential to affect the exercise of human agency**, or humans' capacity to understand a system's functions and its effects in a relevant context; deliberate and decide upon suitable actions in a timely manner; and act in a way where responsibility is guaranteed.<sup>3</sup>

Dynamics of human-machine interaction result in **distributed agency between humans and AI systems, where they are not separated into two distinct entities but rather form part of a socio-technical system**.<sup>4</sup> As part of this system, both sides may influence each other in different ways, which then translate into various forms of distributed agency located along a spectrum. In some instances, dynamics of human-machine interaction will offer more opportunities for exercising human agency in targeting decisions. In other instances, however, the humans involved in use-of-force decision-making will be more constrained in their ability to exercise agency.

For example, **humans' ability to exercise agency might be limited by cognitive biases such as automation bias or anchoring bias**. Humans could over-trust AI DSS even when knowing that there might be malfunctions or unintended errors involved, risking an overreliance on algorithmic outputs without engaging in the critical deliberations and assessments that are needed to exercise human agency, especially in critical targeting decisions that might inflict death, destruction, and severe harm. Such biases are typically exacerbated by the increased speed of AI-assisted military decision-making, especially in contexts where there are high levels of pressure to act rapidly. They can also be exacerbated by AI DSS that are used for prescription or recommendations, because such systems restrict the options or courses of action available to human decision makers.

Moreover, given that AI DSS are likely to be employed not individually but rather as part of a network of systems, the increased complexity of interactions can result in situations where humans act upon some outputs suggested by AI DSS, but do not overall exercise a high quality of agency. Due to these and many other concerns related to interactions between humans and AI DSS, **there is a need to further investigate challenges of human-machine interaction that result in AI DSS not positively 'supporting' humans but rather undermining humans' ability to exercise agency**.<sup>5</sup>

<sup>3</sup> Anna Nadibaidze, Ingvald Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems: A Review of Developments and Debates* (Odense: Center for War Studies, 2024), <https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/>.

<sup>4</sup> Ingvald Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025), <https://www.cigionline.org/publications/human-machine-interaction-and-human-agency-in-the-military-domain/>.

<sup>5</sup> Anna Nadibaidze, "Do AI Decision Support Systems 'Support' Humans in Military Decision-Making on the Use of Force?" *Opinio Juris*, November 29, 2024, <https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force/>.

The risks of not addressing challenges of distributed agency are substantial. First, situations where humans are restricted in their exercise of agency **raise questions about compliance with international humanitarian law**, which requires that humans be held accountable and legally responsible for violations of legal principles. Although humans remain officially in control of the selection and engagement of targets, there are concerns about the exact role played by humans in context of using AI DSS in practice.

Second, these concerns also extend to the risk of **negatively affecting moral agency and responsibility in warfare**. Challenges of human-machine interaction that result in distributed agency would allow humans to feel less morally responsible for decisions that could affect other people's lives. They also risk making the human role a nominal, 'box-checking' exercise which can *de facto* be compared with AI DSS playing an 'autonomous' role because the human role is substantially reduced.

Third, there are **security and operational risks related to distributed agency dynamics**, especially when they give too prominent roles to AI DSS and algorithmic outputs. AI systems often malfunction, are trained on biased sets of data which do not apply beyond the training context or specific contexts of use, as well as integrate assumptions that might not be strategically or operationally beneficial.

Various types of biases, issues of trust, uncertainties, targeting and military doctrines, political and societal contexts in which AI DSS are used – all these aspects **can lead to dynamics of distributed agency which limit the exercise of human agency and prioritize algorithmic outputs**. It is important to investigate these dynamics and ensure that distributed agency provides more opportunities than limitations to human decision makers in warfare.

#### **Relative under-development of the international debate on AI DSS**

Despite increasing reports about the use of AI DSS in recent and ongoing armed conflicts, and the significant challenges and risks they pose to the effective exercise of human agency, **the international debate on human-machine interaction in the use of AI DSS remains insufficiently developed**, particularly within intergovernmental UN settings. Current discussions on AI in the military domain, including those within the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS), have focused on the use of AI at the tail-end of the targeting process, specifically autonomy and AI in weapon systems. This narrow focus risks overlooking or failing to address critical normative, legal, ethical, security, and operational risks that can proliferate and compound throughout the entire targeting decision-making process.

An increasing, albeit still limited, number of stakeholders are raising this issue at international multistakeholder forums, such as the Summits on Responsible Artificial Intelligence in the Military Domain (REAIM). Some international non-governmental organisations and research institutes – such as the International Committee of the Red Cross (ICRC), the Stockholm International Peace Research Institute (SIPRI), the UN Institute for Disarmament Affairs (UNIDIR), and the Asser Institute – have initiated discussions on challenges posed by AI in the military domain, beyond the issue of autonomy in weapon systems. Despite this progress, there remains **a clear need to develop a more comprehensive and inclusive international multistakeholder debate to guide the responsible development and deployment of AI DSS in military contexts**.

## Way forward

In closing, we sketch three ways intended to move the international debate about applications of AI in the military domain forward:

1. **Increase awareness for the implications of practices of designing, developing, and using AI DSS.** States and other stakeholders across industry, civil society, and academia engaged in the governance, development, and use of AI DSS for military targeting must consider the implications of their practices. These practices influence what counts as ‘appropriate’ ways of considering and employing AI DSS and thereby shape what becomes the accepted, requisite quality of human oversight and agency exercised over the whole process of use-of-force decision-making. To increase such awareness, the debate pursuant to [A/RES/79/239](#) at the UNGA First Committee should centrally focus on the issue of AI DSS in the military domain.

2. **Consistently map both ‘best’ practices and ‘problematic’ practices associated with the design, development, and use of AI DSS.** To get a better sense of the direction that the design, development, and use of AI DSS take, states and other stakeholders need to closely map their own (and others’) practices. While there have been some limited efforts to exchange potential best practices, we also need to be attentive to practices with potentially problematic effects. This should encompass practices exercised across the full life cycle of AI systems from development to use and post-use review. Mapping such practices would offer stakeholders a better overview of which practices may be beneficial, i.e., provide opportunities for the exercise of human agency, and which practices may be problematic, i.e., limit the exercise of human agency, and therefore assess the desirability of particular practices.

3. **Pursue the debate on AI DSS within a multistakeholder format.** States should work with diverse stakeholders – including academics across social sciences and technical disciplines, civil society representatives, and international organizations – to develop normative guidance and regulation, especially regarding the human role in military decision-making. Moreover, top-down processes towards governing AI DSS should be accompanied by a bottom-up, standard-setting process focused on establishing operational standards. Such an inclusive approach could strike a balance between national security and humanitarian concerns, while reinforcing the need to ensure that humans can exercise agency in use-of-force decisions.

## Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain

[11 April 2025]

### 1. Introduction

The Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain (GC REAIM) welcomes the opportunity to contribute to the United Nations Secretary-General’s report pursuant to resolution [A/RES/79/239](#).

GC REAIM recognises that military applications of AI present both opportunities and challenges for global peace and security. Accordingly, the establishment of responsible and ethical governance – consistent with States’ obligations under applicable international law – is essential. The global community must take proactive steps to ensure that military AI is developed and deployed in a manner that de-escalates rather than escalates conflicts; respects and enhances, rather than compromises, the sovereignty and territorial integrity of states; promotes rather than threatens the security and safety of civilians; constrains and supports rather than erodes the existing rules-based international order.

In line with GC REAIM's resolute commitment to advancing international governance efforts, this note outlines some of the – non-exhaustive – views expressed by GC REAIM Commissioners and Experts on the implications of AI in the military domain to peace and security. The views presented are general in nature and will be further elaborated in the forthcoming GC REAIM report. While the Commission plans to present substantive and actionable recommendations for stakeholders in September 2025, this note does not yet include concrete proposals. As discussions among Commissioners and Experts are still ongoing, it instead highlights some of the key opportunities, challenges, benefits and risks posed by AI in the military domain to peace and security.

## **2. Technological Foundations**

GC REAIM holds that meaningful policy deliberations on AI in the military domain must be grounded in a shared, foundational understanding of the underlying technologies and their potential trajectories. The complexity of AI technologies often gives rise to misunderstandings, inflated expectations, or misguided applications. Consequently, it is imperative to demystify AI through formal and well-defined frameworks that distinguish between current capabilities and speculative future developments. To support this objective, GC REAIM is developing a taxonomy which seeks to map the full spectrum of AI applications across military and broader peace and security contexts. The taxonomy differentiates between the implications of AI in operational activities – such as warfighting and intelligence – and administrative activities – such as logistics and personnel training and helps identify the specific applications of AI that should be prioritised in governance deliberations.

In its approach to the creation of a taxonomy, GC REAIM highlights the need for and contributes to a concerted effort to clarify, standardise, and encourage the accurate use of technical language with different layers of abstraction for policymakers, experts, and the public, thereby enhancing transparency, mutual understanding, and public trust. GC REAIM also cautions against the uncritical multiplication or adoption of new terminologies in AI governance discourse, unless these are clearly defined; and to ensure such terms are not used to circumvent or obscure existing legal obligations. Precision and consistency in language are the basis of responsible AI governance.

## **3. Implications for Peace, Security, and Stability**

GC REAIM recognises that the integration of AI into the military domain presents benefits as well as both foreseeable and unforeseeable risks to international peace and security. A balanced approach to the range of opportunities and challenges emerging throughout the AI life cycle lies at the core of GC REAIM's method and is essential for responsible AI governance.

AI in the military domain may contribute to international peace and security in several important ways. At the developmental stage, the advancement of military AI capabilities may act as a deterrent to violence, as the mere development and presence of advanced technologies by responsible actors can encourage restraint by aggressors. Military AI may enhance early warning systems, strengthening conflict prevention strategies, and supporting arms control verification through AI-driven tools that foster transparency, trust, and cooperation among states – fundamental elements in conflict prevention. AI can also bolster national security and defence by improving the precision, accuracy, and efficiency of intelligence analysis and situational awareness, enabling real-time threat detection, and facilitating more efficient counterterrorism operations through predictive analytics and autonomous systems. AI-powered systems can rapidly process vast amounts of complex data, enabling military forces



to make timely, informed decisions that may prevent escalation and support conflict de-escalation efforts. These traits can also help improve targeting accuracy and precision, potentially reducing the risk of collateral damage or fratricide – attacks on one’s own forces – and aiding compliance with International Humanitarian Law (IHL) to protect the security of protected persons, such as civilians and non-combatants, during armed conflict. Military AI may also reduce certain forms of human bias and enhance accountability by providing precise data, surveillance, and real-time monitoring, enabling clear attribution of actions to specific actors. In these ways, AI offers meaningful opportunities to reinforce adherence to international law and ethical standards, strengthening the normative foundation of the rules-based international order underpinning global peace and security.

AI in the military domain also presents a range of risks. As with the development of other general-purpose technologies, the development of AI in the military domain may accelerate arms races. AI technologies driven by the commercial market may be repurposed by militaries or soldiers in need or increase the access of violent non-state actors to AI-enabled military capabilities, which may intensify ongoing conflicts and contribute to broader instability. There are also concerns that states could employ AI technologies to suppress human rights, entrench internal repression, and destabilise both regional and global peace.

Concurrently, as with AI more broadly, the environmental consequences of military AI – such as the energy-intensive demands of AI systems, resource extraction, and ecological damage from AI-enabled military systems – could aggravate resource scarcity and environmental degradation, fuelling tensions and undermining long-term peace. However, given the impact militaries have on civilian technology development, efforts to reduce the environmental impact of AI in defence settings could have far-reaching beneficial consequences for all uses of AI. As such, considerations of environmental impacts should be a component of responsible AI governance in the military domain.

The large-scale data extraction required for AI development could intensify geopolitical rivalries, facilitate intrusive surveillance, and create distrust through opaque and exploitative data practices. Such deployment of military AI may perpetuate discrimination and exacerbate social divisions, undermining stability and ultimately international peace and security.

There are simultaneously significant concerns regarding the potential of integration of AI within the command, control, and communication (C3) structures of nuclear weapons. A number of Commissioners and Experts have emphasised that this is a red line that must not be crossed. The commitment of several nuclear-armed states to human decision-making surrounding the employment of nuclear weapons is therefore applauded. Further, the development of large-scale lethal autonomous weapon systems – such as swarms of anti-personnel devices – risks creating a new category of weapons of mass destruction, posing serious threats to global peace and security. Relatedly, AI may lower the barriers to creation and use of nuclear, chemical, or biological weapons by state or non-state actors, thus generating new challenges for arms control and non-proliferation regimes.

Beyond these strategic risks, AI may affect the character of war and lower the thresholds for armed conflict. By increasing the speed of armed escalation and driving changes in the capabilities of weapons systems, AI in the military domain may reduce states’ confidence in their deterrent capabilities – particularly in the face of cyber infiltration risk – thus influencing how decision makers receive, process, and act on information. AI in the military domain could also exacerbate asymmetric warfare and violence by widening technological disparities that could increase the likelihood of force being used prematurely or disproportionately.

Operationally, inaccurate AI systems used for targeting can undermine the security of protected persons under IHL by increasing the risk of indiscriminate attacks, violations of proportionality, and failure to distinguish between combatants and civilians. Closely related to this is the risk of fratricide due to potential errors in target identification or decision-making, which can undermine operational effectiveness, escalate conflict, and erode trust within militaries and alliances. Finally, there are views that the use of certain AI systems in the military domain can create accountability gaps absent clear rules. By complicating the attribution of responsibility for unlawful actions, the deployment of AI in the military domain could undermine key principles of international law and state responsibility for internationally wrongful acts. This may complicate efforts to hold individuals or states responsible for violations, leading to a reduced deterrent effect against unlawful conduct. Without avenues to hold actors legally responsible, the enforcement of international law weakens, potentially destabilising peace, encouraging impunity, and exacerbating global insecurity.

#### **4. Decision-Making and Responsibility**

GC REAIM acknowledges the ethical and legal challenges that arise from integrating AI into military decision-making which may have a direct impact on preservation of peace and security. The relationship between human judgment and machine outputs is complex and without measures to ensure lawful, responsible and effective development and deployment, there can be an erosion of accountability and increased risks of unintended harm. As AI systems become more sophisticated and integrated within military capabilities, it is plausible that algorithmic decisions may become more commonplace across global battlefields, introducing moral and legal challenges regarding human control, oversight and judgment in diverse contexts.

To address these risks, GC REAIM promotes the need for context-appropriate human judgement over specific uses, capabilities and decisions of AI in military applications. The GC REAIM report will list considerations and conditions that underpin and support human responsibility, judgment and means of adequately evaluating relevant actions and decisions. This could include the introduction of technical standards for explainability, as well as maintaining appropriate human oversight in targeting decisions, assessments of precautions, proportionality and distinction, and other critical operational choices. However, given that the very definition of autonomy in machines suggests the minimisation or removal of the human, ensuring human responsibility and accountability may require focusing on human decision-making at earlier stages of a system's life cycle, as the systems structure the behaviour of all who work with it. Human oversight is essential to uphold state obligations under applicable international law, in particular, IHL.

Military AI systems must be designed not only to support all individual and collective agents in the military domain to be effective in safely carrying out their lawful tasks, but also to do so responsibly and without compromising or undermining their status as moral human agents. GC REAIM suggests that military AI based socio-technical systems need to be explicitly and demonstrably designed to adequately attribute and apportion responsibilities and is determined to contribute to this process. For the security of protected persons, parties to armed conflicts should at all times be able to demonstrate that everything possible has been undertaken to create the conditions under which military personnel can effectively apply extant and widely shared principles and laws of armed conflict to their own situation, when using or relying upon AI components in the execution of their tasks.

#### **5. Governance and Regulation**

In light of both the opportunities and risks associated with military AI, GC REAIM supports a comprehensive governance framework that implements authentic international law. GC REAIM reiterates that existing legal regimes provide a solid foundation for regulating AI technologies. Governance must incorporate and account for procedural safeguards (due diligence and legal reviews, transparency of testing, evaluation, and validation, accreditation, and verification), substantive obligations drawn from various branches of international law, and soft law tools (military doctrines, national policies and strategies, norms and standards). In principle, all relevant international legal frameworks must be considered and applied. These include, but are not limited to, the following: (1) international law (*jus ad bellum*) which regulates when and how states use force, codifying a general prohibition on the use of force and exceptions such as in the case of self-defence, (2) international humanitarian law (*jus in bello*) which governs conduct during armed conflict and ensures the security of protected persons, (3) international human rights law.

GC REAIM further emphasises the critical role of international, regional, and domestic institutions in implementing and enforcing these legal norms. Effective governance requires collaboration across these levels and the inclusion of both binding (hard law) and non-binding (soft law) instruments. Soft law mechanisms, such as codes of conduct and ethical principles, can complement existing treaties and facilitate rapid, flexible responses to technological developments.

To address the diverse range of challenges surrounding the integration of AI into the military domain, GC REAIM supports proactive risk-mitigation and confidence-building measures. While binding regimes are challenging for general purpose technologies, there may be opportunities for rigorous monitoring, verification, and enforcement mechanisms inspired by successful global arms control regimes. For example, Commissioners and Experts have discussed ideas such as an Autonomous Incidents Agreement to reduce the risks of miscalculation among AI-enabled autonomous systems, or a committee or consortium that could set guidelines and recommendations surrounding the testing and evaluation of AI systems, including generative AI. GC REAIM also suggests that states and industries should consider adopting human-centred safety-by-design principles, implement red-teaming practices throughout AI system life cycles, and maintain clear chains of accountability for all actors. Only through robust multilateral dialogue and inclusive multi-stakeholder cooperation can AI be effectively governed to enhance peace and security rather than exacerbate global instability.

GC REAIM acknowledges that the development of a comprehensive governance framework for military AI faces several key challenges. First, there is the challenge of diverse interests and perspectives, with states, private companies, and civil society holding varying and sometimes conflicting views on the regulation of military AI. Second, the sensitivity surrounding national security and defence poses a significant barrier, as many states are reluctant to subject their military technologies to international scrutiny or regulation due to legitimate security interests. Third, achieving meaningful and substantive inclusivity in discussions is often difficult, as key stakeholders may be excluded or marginalised in decision-making processes. Fourth, a trust deficit between states, international organisations, and the private sector complicates efforts to establish cooperative governance. Fifth, the presence of crosstalk, incommensurability, and discursive dissonance arises due to the diverse backgrounds and expertise of stakeholders, making consensus-building challenging. Finally, these obstacles are compounded by the lack of clear frameworks that address the complex ethical, legal, and technical issues at the nexus of AI and the military domain. In light of these challenges, the final GC REAIM report will offer strategies to navigate and overcome these barriers in developing a robust governance framework.

## 6. Conclusion

GC REAIM observes that the rapid advancement and deployment of AI technologies in military contexts poses opportunities, challenges, benefits and risks for global peace and security. Balancing these considerations must be met with a technologically sound, inclusive, principled, and legally grounded approach to governance.

A clear understanding of AI's technological foundations is necessary to properly address its role in modern warfare. Ethical and legal responsibility should remain human-centred, and governance frameworks must rely on the robust application of international law, supplemented by cooperative multilateral efforts and soft law instruments when appropriate. In its formation and deliberations, GC REAIM has had the opportunity to reflect upon the conversations happening in broader governance processes, finding ways to effectively bridge gaps between disciplines and regional perspectives.

GC REAIM urges the United Nations and all State Parties to place these principles at the heart of global discussions on the implications of AI in the military and broader peace and security, for the present and future generations. Only through concerted international cooperation, guided by a shared commitment to human dignity, peace, and justice, can we ensure that the future of AI in the military domain is one that strengthens our common security.

## InterAgency Institute

[11 April 2025]

The InterAgency Institute was established in December 2020 as a digital think tank, founded by expatriate and Global South women as a collective of researchers. It is in this condition that we address this submission on “opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems,” following [A/RES/79/239](#). With this, we seek to craft a complementary set of suggestions to develop the policy discussion in points where understanding that AI encompasses a wide array of data-processing techniques, and may be integrated into different types of warfare, in multiple parts of the organization, and at different levels.

The InterAgency Institute would like to point to overarching trends that fall within our areas of expertise, namely: (1) a focus on the global south, specially in how to prevent furthering the security gap; and (2) in how interagency cooperation in a time of greater mistrust may be leveraged to ensure the integration of AI in the military does not. Additionally, we make the point that Decision Support Systems (DSS) create analogous problems when compared to Autonomous Weapon Systems (AWS).

### 1. Addressing the security gap between the Global South and the Global North

The increasing technological intensity and digitalization of the battlefield are likely to increase the capacity gap between countries in the Global North & South. The “optimization of war” entails furthering this discrepancy, augmenting threats, and deteriorating the global security landscape. The wide range of AI-enabled solutions represents discrepant utility levels across tools.

While some tools require a low threshold (thus providing usually an equally low ceiling), the systems that pose the biggest military advantage require a high

knowledge threshold to be implemented, therefore, will likely not be open source, and will only be available to entities with sufficient means to develop or acquire them. Given the experience in past decades on multilateral forums, it is important to recognize that interest in access to these technologies will play a role in the negotiations.

In the long term, the current trend of “technological sovereignty” (or more specifically of restricted technological access due to global inequalities) may be transformed to undermine such technological control, creating far-reaching implications of this new revolution in warfare, involving stakeholders that may be reluctant to shape modern discussions due to a lack of current development of these technologies in their ecosystem.

## **2. InterAgency cooperation in times of distrust**

These issues call for interagency cooperation at both the strategic and operational levels. The lack of interagency cooperation might lead to threat escalation and the eroding foundations for peace and security. Interagency cooperation should focus on formalizing specific channels for communication between different States, developing strategies for AI implementation that will not damage diplomatic relations, and generating more transparency in the interactions between agencies and contractors. The participation of different branches of government at the UN-level discussions is pivotal for a whole-of-government perspective in the deliberations. Beyond interagency cooperation at the governmental level, the wide array of applications of military AI calls for different sets of Confidence Building Measures (CBMs).

Since AI may be integrated in different warfare types and at different levels, its applications for different contexts have different ethical implications and consequences. Therefore, a monolithic understanding of risks posed by AI in the military context and consequently a unique set of CBMs would be inadvisable. CBMs for AI use in the strategic level of cyberspace will not be the same as CBMs for AI use in the tactical level of aerial warfare. Therefore, thinking about CBMs for military AI as a monolith will lead to inaccurate and in some cases inapplicable measures, undermining its effectiveness.

There is a necessity for sharing best practices in the introduction of AI into military procedures. In this sense, a trade-off should be made, prioritizing best practices that contribute to strengthening the aforementioned points of interagency cooperation and CBMs, and other practices that fall within the larger umbrella of strengthening international peace and security. Sharing of best practices relating to cybersecurity and reliability of the technology could also take place, but they should give priority to CBMs that focus on integration of AI at the strategic level and in manners that avoid the escalation of threats.

## **3. Decision Support Systems**

Target identification or recognition via AI-enabled Decision Support Systems (DSS) entail analogue problems to Autonomous Weapon Systems (AWS). Digital dehumanization, lowering the threshold of violence, and automation bias are byproducts of that process that may only be avoided by the creation of red lines prohibiting such systems that replicate those concerns.

This problem stems not only from AI, but from a wider trend. Other data processing techniques that involve deterministic sorting of data that is not adequately processed by human operators also generate these problems. This caveat should be

made to understand that not only systems with AI-enabled technology in DSS pose these kinds of threats, but a wider array of data gathering/processing techniques.

### **Conclusions and recommendations**

- Formal interagency bodies to interface with multilateral AI/military tech negotiations
- Funding and support for academic research in the Global South focused on military AI implications;
- Regular technical-diplomatic summits focused on transparency, shared definitions, and threat perception;
- Prioritize capacity-building initiatives for Global South actors;
- Red lines and confidence building measures could be tailored to the specific technology and operational context;
- The discussions on Autonomous Weapon Systems encapsulate worries around AI-enabled Decision Support Systems. The creation of red-lines for these systems could benefit from building upon recommendations of the GGE on LAWS;

### **International Committee for Robot Arms Control**

[11 April 2025]

The International Committee for Robot Arms Control (ICRAC) values the opportunity to submit our views to the United Nations Secretary-General in response to Resolution [A/RES/79/239](#) “Artificial intelligence in the military domain and its implications for international peace and security.”

Founded in 2009, ICRAC is a civil society organization of experts in artificial intelligence, robotics, philosophy, international relations, human security, arms control, and international law. We are deeply concerned about the pressing dangers posed by AI in the military domain. As members of the Stop Killer Robots Campaign, ICRAC fully endorses their submission to this report, and wishes to provide further detail regarding the concerns raised by AI-enabled targeting.

Increasing investments in AI-based systems for military applications, specifically AI-enabled targeting, present new threats to peace and security and underscore the urgent need for effective governance. ICRAC identifies the following concerns in the case of AI-enabled targeting:

1. AI-enabled targeting systems are only as valid as the data and models that inform them. ‘Training’ data for targeting requires the classification of persons and associated objects (buildings, vehicles) or ‘patterns of life’ (activities) based on digital traces coded according to vaguely specified categories of threat, e.g. ‘operatives’ or ‘affiliates’ of groups designated as combatants. Often the boundary of the target group is itself poorly defined. Although this casts into question the validity of input data and associated models, there is little accountability and no transparency regarding the bases for target nominations or for target identification. AI-enabled systems thus threaten to undermine the Principle of Distinction, even as they claim to provide greater accuracy.

2. Human Rights Watch research indicates that in the case of IDF operations in Gaza, AI-enabled targeting tools rely on ongoing and systematic Israeli surveillance of all Palestinian residents of Gaza, including with data collected prior to the

current hostilities in a manner that is incompatible with international human rights law.

3. The increasing reliance on profiling required by AI-enabled targeting furthers a shift from the recognition of persons and objects identified as legitimate targets by their observable disposition as an imminent military threat, to the ‘discovery’ of threats through mass surveillance, based on statistical speculation, suspicion and guilt by association.

4. The questionable reliability of prediction based on historical data when applied to dynamically unfolding situations in conflict raises further questions regarding the validity and legality of AI-enabled targeting.

5. The use of AI-enabled targeting to accelerate the scale and speed of target generation further undermines processes for validation of the output of targeting systems by humans, while greatly amplifying the potential for direct and collateral civil harm, as well as diminishing the possibilities for de-escalation of conflict through means other than military action.

Justification for the adoption of AI-enabled targeting is based on the premise that acceleration of target generation is necessary for ‘decision-advantage’, but the relation between speed of targeting and effectiveness in overall military success, or longer-term political outcomes, is questionable at best. The ‘need’ for speed that justifies AI-enabled targeting is based on a circular logic, which perpetuates what has become an arms race to accelerate the automation of warfighting. *Accelerating the speed and scale of target generation effectively renders human judgment impossible or, de facto, meaningless.* The risks to peace and security – especially to human life and dignity – are greatest for operations outside of conventional or clearly defined battlespaces. Insofar as the use of AI-enabled targeting is shown to be contrary to international law, the mandate must be to *not* use AI in targeting.

In this regard, ICRC notes that the above systems present challenges to compliance with various branches of international law such as international humanitarian law (IHL), *jus ad bellum* (UN law on prohibition of use of force), international human rights law (IHRL) and international environmental law. In the context of military AI’s implications for peace and security, *jus ad bellum*, a framework that prohibits aggressive military actions and regulates the conditions under which states may lawfully resort to the use of force, is the most relevant. In the same manner IHRL is important in this context because it is designed to uphold human dignity, equality, and justice – values that form the foundation of peaceful and secure societies.

## International Humanitarian Law and Youth Initiative

[11 April 2025]

Artificial intelligence (AI) has gained a universal recognition during the 1950s’. Technological emergence has assisted humans in almost all facets of their lives thereby making work easier and faster. Moreso, the rapid growth of Artificial intelligence in technological field enthraling commercial investors, law makers, defense intellectuals and international competitors can be evidential in theoretical premises of international security. The use of Artificial intelligence (AI) in modern warfare particularly in the In the Middle East and North Africa, Ukraine/Russian armed conflict which has resulted in the killings of thousands of innocent civilians with women and children being the most vulnerable. The emergence of AI is expected to be utilized in improving all sectors in our daily lives However, its Negative application in the military domain continues to create Humanitarian crisis between

warring parties making it of regional and international concern. The war in Gaza is one of the deadliest and most destructive war in history with technology playing a central role in enabling mass slaughter and destruction ranging from supplying the dystopian systems used to automate the killings and bombing.<sup>1</sup> Following the October 7 2023, there have been extensive reports evidencing the Israeli occupation forces use of surveillance technology, artificial intelligence, and other digital tool to determine who, what and when to attack in Gaza trip. Thus, this violates the principles of international humanitarian law which emphasize the necessity of distinguishing those in active combat and not<sup>2</sup> and to take necessary precautions when conducting an attack to minimize civilian harm.

IHLAI in this paper, responding to the request of the UN Secretary-General pursuant to a resolution [A/RES/79/239](#), adopted by the General assembly on 24 December 2024 on Artificial intelligence in the military domain and its implication to international peace and security therefore, it analyzes AI In modern warfare, its implication to international peace and security and the role of technological companies in armed conflict.

### **Artificial Intelligence in Modern Warfare: A Legal and Humanitarian Perspective**

The rules of international humanitarian law do not explicitly address the use of modern technological tools and artificial intelligence (AI) during armed conflicts. However, its core principles – such as distinction, proportionality, and precaution – remain applicable and binding on all parties. These principles require the differentiation between military objectives and civilians, and oblige parties to take all feasible measures to avoid or minimize harm to civilian populations. In recent years, militaries have contracted private companies to develop autonomous weapons systems. However, the armed conflict in Gaza stands out as one of the most prominent cases where commercially developed AI models – originally created in countries like the United States – have been employed in actual combat operations, despite the fact that these systems were not initially designed to make life-or-death decisions.

This shift highlights a troubling rise in the militarization of technology without clear legal or ethical oversight. While some of these tools may enhance operational efficiency, their unregulated use poses serious risks of human rights violations, especially amid a lack of transparency about how these tools function, the origin of the data they rely on, and the accuracy of their outcomes<sup>3</sup>.

One of the most pressing concerns recently raised is the deployment of digital military tools based on unreliable data or flawed algorithms. Some of these systems depend on mass surveillance of Gaza's<sup>4</sup> population, including the collection of personal data prior to the outbreak of hostilities. Such practices raise legal and ethical questions regarding their compatibility with international obligations to safeguard privacy and prevent the misuse of personal information for the purpose of direct targeting.

Among the tools reportedly in use is a system that tracks population movement through mobile phone data to monitor evacuations from certain areas. Another generates lists of structural targets to be hit militarily. A third tool classifies

<sup>1</sup> Accessnow. (October 2024) Big Tech and the risk of genocide in Gaza: what are companies doing? Available at <https://www.accessnow.org/gaza-genocide-big-tech/>.

<sup>2</sup> Article 48 of Additional protocol I of the Geneva convention.

<sup>3</sup> Human Rights Watch, "Israel: AI-Powered Targeting Systems May Be Committing War Crimes in Gaza", 2024.

<sup>4</sup> Associated Press, "Documents Reveal Israel's Use of AI Tools in Targeting Gaza", Investigative Report, 2024.



individuals based on levels of suspicion regarding their affiliation with armed groups, while a fourth seeks to determine the precise location of a target in order to carry out a strike at the opportune moment. These tools largely rely on data extracted from mobile devices – whether through cell tower location information or GPS<sup>5</sup>. However, from a technical perspective, such data is insufficiently precise to confirm an individual's presence at a specific location at a given time, particularly in conflict zones where individuals frequently change phones or numbers. Over-reliance on this technology may lead to fatal mistakes, especially when a mobile phone is used as a substitute for verifying a person's actual presence in a targeted area. Legally, the use of such systems without taking all feasible precautions to protect civilians constitutes a clear violation of international humanitarian law – particularly Article 57<sup>6</sup> of Additional Protocol I to the Geneva Conventions, which obliges parties to take constant care to spare civilian lives during military operations.

Given this reality, urgent questions must be raised about the future of AI in warfare and the legislative and legal mechanisms needed to regulate it. Without proper oversight, these tools risk becoming instruments of systematic human rights abuses rather than technologies aimed at ensuring greater protection for those affected by war.

### **Implications of Artificial Intelligence on International Peace and Security**

Armed conflicts in various regions around the world, such as Gaza, Lebanon, Syria, Ukraine, and Libya, have had catastrophic humanitarian and security consequences. These conflicts have led to the mass displacement of civilian populations, depriving thousands of people of their basic rights such as food, water, shelter, and healthcare. These individuals live in dire humanitarian conditions, with a significant increase in deaths due to famine, thirst, and diseases caused by contaminated water, in addition to exposure to harsh weather conditions without protection.

In this context, the increasing use of artificial intelligence and drones as weapons in conflicts, particularly by Israel in the Gaza Strip<sup>7</sup>, stands out. Since October 2023, there has been a notable escalation in the use of “quadcopters” to carry out precise and targeted strikes against civilians. These drones are equipped with data analysis algorithms and offensive capabilities, enabling them to target individuals based on tracking their movements or mobile phone signals.

According to documented reports, this technology has led to the death of more than 1,000 Palestinians by May 2024, including a significant number of women and children. This constitutes a grave violation of international humanitarian law, particularly Articles 51 and 57 of Additional Protocol I to the Geneva Conventions, which prohibit attacks on civilians and obligate parties to the conflict to take all necessary precautions to avoid harming them.

The concerns are not limited to the use of artificial intelligence against individuals but extend to the misuse of data. Relying on mobile phone tracking technologies (either through GPS data or cell tower signals) to pinpoint individuals' locations presents serious risks. Recent studies have shown that these systems do not provide enough accuracy to reliably determine someone's location, especially in conflict zones where phones may be

<sup>5</sup> Human Rights Watch (2024). Questions and Answers: Israeli Military's Use of Digital Tools in Gaza Available at Questions and Answers: Israeli Military's Use of Digital Tools in Gaza | Human Rights Watch.

<sup>6</sup> Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), Article 57.

<sup>7</sup> TRTWORLD (2024) Quadcopter strikes: 1000 Palestinians killed by Israeli drones in one year. Available at Quadcopter strikes: 1000 Palestinians killed with drones in a year.

swapped or disconnected frequently. This means that relying on these methods without field verification can lead to erroneous decisions, resulting in unlawful killings.

In a well-known case, a Palestinian woman named “Silah” was killed while carrying a white flag and leading her family to safety. After stepping onto a main street, she was targeted by a small drone that shot her in the head. This incident, witnessed by those around her, serves as a stark example of the disastrous outcomes of unregulated use of technology on the battlefield<sup>8</sup>.

In Libya, drones played a decisive role in the battles between conflicting parties, particularly as many of these drones, including Turkish and Chinese models, were operated using data analysis systems to target objectives. Some of these systems are believed to rely on artificial intelligence techniques for targeting, without legal oversight. The use of these tools in urban areas like Tripoli and Sirte has led to the deaths of civilians and extensive damage to infrastructure<sup>9</sup>.

All of these events indicate that integrating artificial intelligence into managing and directing armed conflicts without an internationally binding legal framework to regulate its use could open the door to widespread violations, especially if these systems are not subject to independent and transparent oversight to ensure compliance with international humanitarian law and human rights.

### **Roles of Companies Developing AI in Armed Conflicts**

Through a rapid increase in artificial intelligence and computer services, U.S. tech corporations have discreetly given Israel the ability to monitor and kill many more militants in Gaza and Lebanon more quickly. However, the death toll among civilians has also skyrocketed, raising concerns that these instruments may be causing the deaths of innocent people. Israel’s recent wars are a leading example of commercial AI models developed in the United States being used in active warfare, despite concerns that they were not originally designed to help decide who lives and who dies.

For years, militaries have hired private companies to create customized autonomous weapons. Numerous American software companies have backed Israel’s battles in recent years, including Microsoft and the San Francisco-based startup OpenAI. Under “Project Nimbus,” a \$1.2 billion contract signed in 2021<sup>10</sup> when Israel first tried out its in-house AI-powered targeting systems, Google and Amazon offer cloud computing and artificial intelligence services to the Israeli military. The military has made use of Dell and Cisco data centers and server farms. Palantir Technologies, a Microsoft partner in U.S. defense contracts, has a “strategic partnership” that provides AI systems to support Israel’s war efforts, while Red Hat, an independent IBM company, has also supplied cloud computing technologies to the Israeli military.

Furthermore, through a number of programs, Microsoft also supplies Israel’s government with services that have allegedly been used to help the Israeli military, police, Israeli Prison Service (IPS), and illegal settlement operations. Over 10,000 Palestinians are being held by the IPS as of October 2024; half of them have been

<sup>8</sup> Gaza grandmother gunned down by Israeli sniper as child waved white flag," *Times Kuwait*, November 2024, <https://timeskuwait.com/news/gaza-grandmother-gunned-down-by-israeli-sniper-as-child-waved-white-flag>.

<sup>9</sup> France 24. (2021). “Have Killer Drones Been Deployed in Libya?”. France 24. Retrieved from <https://rb.gy/1m6k43>.

<sup>10</sup> APNEWS (2025). As Israel uses US-made AI models in war, concerns arise about tech’s role in who lives and who dies. Available at How US tech giants' AI is changing the face of warfare in Gaza and Lebanon | AP News.

detained without being charged or having a trial date scheduled. At least 310 medical professionals, UN employees, women, and children are among the Palestinian prisoners from Gaza who are presently detained in prolonged, secret, and incommunicado detention, where they are subjected to torture, mistreatment, and sexual violence and abuse, according to the UN Human Rights Office.

Companies are under obligation to respect human rights within their scope of operations. Companies that directly aid the offender – for example, by offering financial, logistical, military, or intelligence support – may be held criminally responsible for a crime committed during an armed conflict. Companies and their managers or executives may be held accountable in certain situations even if they had no direct involvement in the crime or no intention of supporting it. As the Office of the High Commissioner on Human Rights (OHCHR) noted, companies “should treat this risk in the same manner as the risk of involvement in a serious crime, whether or not it is clear that they would be held legally liable<sup>11</sup>.”

In light of the concerns raised in this submission and their implications for international peace and security, IHLYI urges states to:

1. **Refrain from the use of AI in military applications:** States should immediately halt the use of artificial intelligence in military activities and establish national regulations and laws to prevent its deployment in warfare.
2. **Work towards a global ban on the military use of AI:** States should actively pursue international agreements and frameworks to ban the use of AI in military contexts, ensuring that no country utilizes AI for warfare.
3. **Avoid the development of autonomous and AI-enabled weapon systems:** States should refrain from developing autonomous weapon systems or AI-powered weaponry that could be used to target humans, ensuring human oversight and decision-making in military actions.
4. **Ensure the protection of personal data:** States must guarantee that personal data is protected from misuse by military forces, law enforcement agencies, border control, and private contractors collaborating with these entities.
5. **Promote accountability in AI development:** Technology companies, researchers, engineers, and financial institutions should commit to not supporting the development or funding of AI technologies designed for military applications, advocating for responsible innovation in line with humanitarian principles.

## Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand

[21 May 2024]

Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand welcome the opportunity to contribute our views to the UN Secretary-General’s report on artificial intelligence (AI) in the military domain and its implications for international peace and security. Our submission briefly outlines our involvement in this issue, and has three sections summarising our position on: a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed; b) Key focuses of a new international instrument; and c) Scope of a new international instrument. The points below are based on discussions with our member and supporting groups about the content of this submission.

<sup>11</sup> Accessnow (2024) Big Tech and the risk of genocide in Gaza: what are companies doing? <https://www.accessnow.org/gaza-genocide-big-tech/>.

## Introduction

Peace Movement Aotearoa is the national networking peace organisation in Aotearoa New Zealand, established in 1981 and registered as an Incorporated Society in 1982. Our purpose is networking and providing information and resources on peace, humanitarian disarmament, human rights and social issues; and we have extensive national networks of member and supporting groups and individuals. We are a founding member of the Stop Killer Robots campaign and coordinate the national Stop Killer Robots Aotearoa New Zealand (SKRANZ) campaign.

SKRANZ was launched in April 2013 to support the global campaign, with a specific national focus on urging New Zealand to take national action to prohibit the development, production and use of autonomous weapon systems; and to take international action to support negotiations on a new treaty to prohibit autonomy in weapon systems. Since 2023 we have widened our focus to include military use of AI as its perils became increasingly obvious.

### (a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed

As outlined in our submission for the UN Secretary-General's report on autonomous weapon systems ([A/RES/78/241](#)) last year, it has been clear for some years now that rapidly developing technological advances in the use of force and increasing autonomy in weapon systems pose an unprecedented threat both to humanity and to the foundations of international human rights and humanitarian law, which are based on respect for human life and dignity, protection of humanity in times of oppression and armed conflict, and human responsibility and accountability for harm.

The serious ethical, humanitarian, legal, and security concerns posed by these developments have been discussed for more than a decade within United Nations bodies – including the Human Rights Council, meetings related to the Convention on Certain Conventional Weapons and in the UN General Assembly – as well as in regional and national governmental and non-governmental forums.

Even as these discussions have taken place, some states have increasingly incorporated autonomy into military use of force in ways that have already resulted in gross violations of international law with disastrous consequences for civilian populations. It is apparent that the absence of specific international law on autonomy in weapon systems, and with differing interpretation by some states as to how existing law applies to new technological developments, the risk of proliferation of ever more dangerous and uncontrollable weapon systems is increasing rapidly.

The need for urgency for international action on this has been highlighted over the past eighteen months by, for example, Israel's use of AI-powered target suggestion systems in Gaza to make high explosive strikes on numerous targets possible in a short time frame, resulting in indiscriminate slaughter of civilians and systematic destruction of life-sustaining infrastructure. The reality of digital dehumanisation with catastrophic consequences is now very evident, as is the increasing tendency towards the development and use of autonomous weapon systems that will remove any remaining vestige of humanity from war.

We have noted with concern that states who brought forward [A/RES/79/239](#) include states that have armed and supported Israel's genocidal attacks on Gaza, and where big data tech companies contributing data storage and AI capabilities to Israel's military systems are based.

Similarly, 'responsible AI in the military domain' (surely an oxymoron) is being promoted by states already developing their own AI targeting and autonomous

weapon systems, as a way of undermining the push towards a binding instrument to prohibit these critical threats to international peace and security.

The US ‘Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy’ has highlighted for us the risk of horizontal proliferation of both military use of AI and autonomous weapon systems as states that do not have their own capability in this regard move from interoperability to integration with the states of armed forces that do or that are developing it. In the case of New Zealand, for example – as it seeks to be a ‘combat capable force multiplier with enhanced lethality’<sup>12</sup> – this involves closer military integration with the armed forces of Australia and the US: New Zealand endorsed the US ‘Political Declaration’ early last year specifically to be compliant with US military doctrine.

These unfortunate developments underscore the urgent need for a new international instrument on military AI and autonomy in weapon systems to clarify and strengthen existing law. The instrument must include both prohibitions and regulations, as outlined below, and must include military use of AI in combat.

As emphasised in the UN Secretary-General’s 2024 report on autonomous weapon systems<sup>13</sup>, negotiations on a new instrument must begin without any further delay, in a multilateral forum where states can come together to work constructively, where the voices of those whose lives have already been impacted by military use of AI and increasing autonomy in weapon systems can be heard, and where UN agencies, the International Committee of the Red Cross (ICRC), and NGOs are active participants.

**(b) Key focuses of a new international instrument**

While much of the work around military use of AI and autonomous weapon systems has focused on the issue of meaningful human control over the use of force, it is our view that the key underlying ethical imperative is preventing human beings from being targeted or attacked by any system utilising digital code and/or sensors. A prohibition on military use of AI and autonomy in weapons systems that are designed or used to target human beings must be the starting point.

Meaningful human control over the use of force clearly has an ethical component, but it is also a practical and legal means to ensure accountability for any autonomy in weapon systems that breach the key dictates of humanitarian law.

**(c) Scope of a new international instrument**

It is our view that a new international instrument should include overarching rules to establish a framework for evaluating current and future technological developments, while promoting increased compliance with international human rights and humanitarian law.

Such overarching rules would prohibit autonomous weapon systems that are designed or used to target humans, and lay out specific obligations to ensure meaningful human control over other systems: for example, that the human operator/s understand the capabilities and limitations of the system, are able to fully evaluate the context in which the system will be used, and are making mindful firing decisions rather than assuming the technology is accurate – this would act to regulate autonomy in weapon systems. It would be useful to specify that decisions made by states on their assessment of new or altered weapon systems that incorporate autonomous features or functions must be transparent.

Furthermore, in the context of the UN Secretary-General’s forthcoming report on AI in the military domain and in the light of the awful consequences of military

<sup>12</sup> See, for example, the 2025 Defence Capability Plan released this week.

<sup>13</sup> Lethal autonomous weapons systems: Report of the Secretary-General (A/79/88), 1 July 2024.

use of AI in Gaza, the scope of a new international instrument must go beyond autonomous weapon systems. It is very clear that there is a spectrum of harmful military use of autonomy, ranging from target decision support systems (as some have described systems such as Lavender), data-based targeting systems, generation of target lists by algorithm or AI, sensor-based targeting systems, through to weapon systems that combine these elements and incorporate varying degrees of machine learning to make target selection decisions and attack autonomously.

We note the 2023 Joint Call by the UN Secretary-General and ICRC President stated “*The autonomous targeting of humans by machines is a moral line that we must not cross*”<sup>14</sup>, yet that has already happened – a point reiterated in the UN Secretary-General’s 2024 report<sup>15</sup>.

It is therefore our view that a new instrument must cover military use of AI – including systems that automate significant decision-making in the use of force, such as target generation, force deployment, and engagement – as well as autonomous weapon systems.

Finally, although we have referred in this submission to military use of AI and autonomy in weapon systems, prohibitions and regulations in a new international instrument must also apply to all coercive agencies of the state, including those used for policing and internal security, for border control, in corrections facilities and in places of detention.

## **Ploughshares**

[11 April 2024]

Project Ploughshares, a Canadian peace research institute, has for over a decade focused its advocacy and research on the military applications of emerging technologies, including artificial intelligence (AI) and autonomous weapons. As AI systems are rapidly advancing and being tested in contemporary conflict zones, international governance frameworks have struggled to keep pace. Meanwhile, intensifying geopolitical competition increases the likelihood that AI technologies will be deployed in complex, dynamic environments for which they are not suited – raising significant risks for civilians.

The wide-ranging use of AI in military applications demands urgent and coordinated international attention. We encourage the Secretary-General and member states to focus on three particularly pressing areas: the use of AI in decision-support systems related to the use of force, the dual-use nature of AI technologies, and the widening capacity gap among states engaging in multilateral discussions.

### **AI decision-support systems**

One area that remains insufficiently addressed in current international discussions is the use of AI in military decision-making, especially decisions about the use of force. Of particular concern are AI-enabled targeting tools such as “Lavender” and “Gospel,” reportedly used in Gaza. These systems are classified as “decision support” because a human is technically required to approve target selections. However, there is little transparency regarding how these decisions are

---

<sup>14</sup> Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 5 October 2023.

<sup>15</sup> As at note 3.

made, how frequently AI-generated recommendations are rejected, or whether human operators fully understand how the AI systems reach their conclusions.

In practice, these systems raise the risk of “rubber-stamping,” in which human oversight becomes superficial, thereby undermining the principle of meaningful human control and increasing the likelihood of harm to civilians. The potential use of such AI systems in early-warning, surveillance, reconnaissance, and nuclear command-and-control systems further amplifies these concerns.

To mitigate these risks, states must work toward clear norms, regulations, and training requirements that enhance operator understanding, counter automation bias, and ensure genuine human engagement in decision-making processes.

### **Dual-use challenges**

AI’s dual-use nature – its applicability to both civilian and military domains – creates further governance complexity. Civilian-developed technologies can be repurposed for military use without appropriate testing or safeguards, increasing the risk of conflict escalation, misuse, and error. Additionally, the accessibility of certain AI tools means that nonstate armed groups may also gain access, potentially using them to target civilians and infrastructure.

We urge states to develop policy mechanisms, including export controls, technology impact assessments, and multistakeholder engagement, to account for dual-use risks and promote responsible innovation.

### **Capacity- and knowledge-building**

Current multilateral discussions reveal stark capacity disparities among states, many of which do not have the resources or technical expertise to participate meaningfully in governance efforts. To ensure inclusive and equitable global engagement, we recommend that states collaborate with the UN Office for Disarmament Affairs to strengthen capacity-building initiatives.

The scientific and academic communities also have a role to play in supporting the development of accessible resources and training materials. International forums, such as the upcoming REAIM Summit in Spain, should include dedicated sessions for knowledge-sharing, especially to support representatives from under-resourced states.

### **Final thoughts**

The international community is at a crossroads. The accelerating militarization of AI demands robust diplomatic responses. We can – and must – move from aspirational principles to concrete, enforceable frameworks, by employing political will, inclusive dialogue, and cross-sector collaboration.

AI-powered warfare is no longer a theoretical risk; it is a present reality. Whether this new era enhances global security or undermines it will depend on the steps states take now to strengthen governance, manage technological competition, and uphold international humanitarian norms.

Without timely, coordinated action, the risks of accidental escalation and unintended conflict will only increase.

## **Soka Gakkai International**

[10 April 2025]

The Soka Gakkai International (SGI) welcomes the opportunity to share our views on the important issue of artificial intelligence (AI) in the military domain. As an NGO whose work is guided by Buddhist principles, we urge that the United Nations, its Member States and other stakeholders take into careful consideration the impact of AI in the military domain from a standpoint of upholding and respecting human dignity.

### **Introduction**

AI in the military domain is rapidly evolving and transforming modern warfare and international peace and security. These systems are being used for various purposes, including surveillance, autonomous weapons, decision-making support, and logistics. With such wide-ranging applications, the integration of AI technologies in military systems poses significant challenges. To better ensure compliance with international humanitarian law (IHL) and uphold protection for civilians and combatants alike there are several issues that we may consider.

### **Lack of transparency and accountability**

- If an AI system were to make an error – such as identifying a target incorrectly – it could be difficult to pinpoint the cause of the error, “the black box problem”. Was it a flaw in the data used to train the AI, an issue with the algorithm, or a problem in the operational context? Without transparency within these systems, assigning responsibility is difficult.
- International laws and treaties, such as the Geneva Conventions, were created before AI systems became commonplace in warfare. Without global norms and legal frameworks, there is no consistent approach to ensuring accountability for AI decisions made in warfare.
- With inadequate accountability mechanisms in place, AI could be used for military strategies that violate human rights, suppress civil liberties, or engage in unethical operations.

### **Speed of decision-making and risk of escalation**

- The ability of a military force to make decisions and execute actions faster than its opponent is increasingly viewed as having a strategic advantage. However, the drive for speed can lead to unintended and costly consequences.
- Decisions made too quickly without proper analysis or consideration can lead to poor outcomes, including tactical blunders, strategic missteps, or ethical violations.
- Instead of diffusing a tense situation or negotiating, if combatants react too quickly it could provoke an even greater confrontation, further escalation and prolonged conflict resulting in more human suffering including amongst civilians.
- The acceleration of decision-making processes closes down the possibility of meaningful human control, the growing trend to automate decision-making threatens the ability to achieve human oversight which is essential to facilitate compliance with IHL.

### **Bias in AI in the military domain**

- AI bias refers to the presence of systematic and unfair discrimination in AI systems, such as historical bias, where systems may reinforce harmful stereotypes, bias in data processing and algorithm development which can lead to making biased decisions and bias in how the systems are used.



- AI bias in the military domain is a significant concern, particularly as AI systems are increasingly being integrated into defense and security operations. The potential for AI bias to emerge in these areas can result in human rights implications, exacerbating existing inequalities and lead to deadly consequences for certain groups.
- AI heavily relies on vast amounts of high-quality and reliable data for decision-making. There are several potential violations when it comes to obtaining this data including issues around privacy and surveillance, challenges of bias also arise when dealing with incomplete and inaccurate data.
- When AI systems are biased, they not only perpetuate inequalities but also contribute to the digital dehumanization<sup>16</sup> of marginalized groups.

### **Proliferation**

- Nations may rush to develop AI-based military technologies to outpace their adversaries, which could lead to a destabilizing arms race and increased global tensions.
- Without regulation autonomous weapons systems in particular, could proliferate globally, including amongst non-state actors which could increase crime nationally and regionally, exacerbating social inequalities, overwhelm resources and infrastructures of countries, as well as undermine social and national security.

### **Conclusion**

The issue of AI within the military contexts is complex, and without regulation, it could lead to serious consequences for global peace and security. The desire to speed up decision-making processes within this context has yet to be proven as an effective way of resolving conflicts and achieving peace and security. Furthermore, you cannot divorce AI in the military and AI in civil uses, a failure to address AI in a military context could have widespread repercussions in all spheres of civil life including law enforcement, border control, education, housing and health care. Fundamentally, AI is here to stay, how we utilize it in the military and in our lives will shape the course of humanity. We have the possibility and the responsibility to decide how we want to use technology, knowledge, and the world's resources. To use it in a way that uplifts humanity or degrades it? This is an urgent question that requires moral, ethical and courageous leadership.

### **Stop Killer Robots**

[11 April 2025]

The Stop Killer Robots campaign welcomes the opportunity to submit our views to the United Nations Secretary-General in response to Resolution [A/RES/79/239](#).

Established in 2012, we are a coalition of more than 270 non-governmental organisations working across 70 countries.<sup>1</sup> We seek to counter threats to humanity and human dignity through the adoption of a new international treaty to prohibit and regulate autonomous weapons systems.<sup>2</sup> We support the development of legal and

<sup>16</sup> Digital dehumanization is a process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

<sup>1</sup> See [www.stopkillerrobots.org/about-us](http://www.stopkillerrobots.org/about-us) and [www.stopkillerrobots.org/a-global-push/member-organisations](http://www.stopkillerrobots.org/a-global-push/member-organisations).

<sup>2</sup> See <https://www.stopkillerrobots.org/our-policies/>.

other norms that ensure meaningful human control over the use of force, counter digital dehumanisation, and reduce automated harm.<sup>3</sup>

### **Building an effective international response to emerging technologies**

Autonomous weapons systems, ‘AI in the military domain,’ and trends and developments in increasingly automated decision-making and action in the use of force – as well as in our lives and societies more broadly – are all part of the same concerning picture:

The growing influence of computer processing and algorithmic thinking increasingly shapes our interactions in the world and the outcomes available to us. There are clear threats to peace, justice, dignity, human rights, equality, responsibility and accountability, and respect for law. We are getting closer to machine processes determining whom to kill.

To address these challenges effectively, a comprehensive and holistic response is needed from the international community.

Adopting a legally binding instrument on autonomous weapons systems will be one critical component: we must draw basic red lines for humanity against the automation of killing, which brings under jeopardy both international humanitarian law and international human rights law, in particular the presumption of innocence, the right to equality and non-discrimination, dignity, and wipes away contextual circumstances of the target(s) in question. The UN Secretary-General’s comprehensive report last year reiterated his urgent call on states to negotiate a legally binding instrument to prohibit and regulate these systems by 2026.

But, a new international treaty on autonomous weapons systems alone may not be enough. States must also reach agreement on preventing and addressing grave harm from other uses of emerging technologies. A whole set of strong international rules are needed that stop the erosion of meaningful human control and the slide towards greater digital dehumanisation and automated harm, across international and domestic practice, in armed conflict and in civilian life.

### **‘Military applications of AI’ are already contributing to civilian harm**

The risks of integrating AI into the use of force in armed conflict reach far beyond those to peace and security between states: a holistic consideration of peace and security that considers dimensions such as ethical, legal, and humanitarian issues must be taken into account in the UN Secretary-General’s report under resolution [79/239](#).

We are already seeing grave threats to civilian protection and human rights and huge harm being caused by AI and automation in the use of force. This is arising from the quest for speed in warfare, the reduction of people to objects, and issues such as automation bias that Stop Killer Robots has raised the alarm about for years.

We have been horrified by reports of the use of AI-powered ‘decision support systems’ by Israel in Gaza, which suggest human targets to strike.<sup>4</sup> According to reports, human approval of these suggestions in vast volumes at high speed has been minimal – entailing digital dehumanisation, the erosion of meaningful human decision-

---

<sup>3</sup> See [www.stopkillerrobots.org/vision-and-values/](http://www.stopkillerrobots.org/vision-and-values/).

<sup>4</sup> ‘Lavender’: The AI machine directing Israel’s bombing spree in Gaza, +972 Magazine <https://www.972mag.com/lavender-ai-israeli-army-gaza/>.

making and control (including through automation bias), and directly contributing to massive and devastating harm to civilians in Gaza, alongside other tools.<sup>5</sup>

Many other states are developing and using such ‘decision support systems’, which raise concerns around international humanitarian law, human rights law, and ethics. So far there are few reports on how these are being deployed, with what constraints and with what impacts. Nevertheless, the push by many states to develop and integrate AI and autonomy into decision-making and the use of force is a huge concern. The further use in hostilities of these kinds of tools by any state in the unacceptable ways that we have seen in Gaza must be prevented. Stop Killer Robots struggles to see how such uses could meet the definition of the responsible application of AI in the military domain given in resolution 79/239.

### **Further risks to peace and security, rights, and human dignity**

The quest for greater speed through AI and automation – towards the goal of increasing the tempo of conflict to a point beyond human cognition in the pursuit of a military and strategic edge – is an extremely dangerous one for international peace and security. These risks are further to the impact ‘AI in the military domain’ is already having on civilian protection. Risks include unwanted escalation, lowered political thresholds to the use of force, and arms race dynamics.

Technologies that can contribute to target selection (such as threat detection tools) and remote biometric surveillance (such as facial recognition) have already had documented negative impacts on human rights such as the rights to privacy, equality and non-discrimination, freedom of expression and peaceful assembly, and the freedom of movement. In the case of facial recognition for identification (1:n), the technology is considered by many legal experts as wholly incompatible with international human rights law.

That AI systems inevitably encode and reproduce the biases of our societies – including racism, sexism and ableism – and that such bias cannot be eliminated, is also well established. The use of such systems to process people in the use of force will inevitably lead to disproportionate – and multiplied – impacts on already marginalised and minoritised people. Integrating automation and AI into decisions and actions in the use of force against people contributes to digital dehumanisation – the process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

### **The relationship with autonomous weapons systems**

Stop Killer Robots notes that the UN Secretary-General’s report will be on the “application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems.” It is important nevertheless to highlight that various applications beyond the boundary of autonomous weapons systems are closely linked to them.

Firstly, such tools could be integrated as components of autonomous weapons systems now or in the future. For example, a ‘decision support system’ could be used as an autonomous targeting system, connected to a platform tasked to strike targets on the list generated, based on processing sensor data. Secondly, these tools are linked not only practically, but raise and are part of the same picture of concern. Strikes undertaken

<sup>5</sup> Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza, Human Rights Watch, <https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza>.

based on the nominal human approval of targets generated by a decision support system do not sit far from strikes undertaken with an autonomous weapons system.

It is therefore important that states consider these issues in dialogue: many of the rules and principles developed for autonomous weapons systems on keeping control and rejecting automated killing will need to be extended (with adaptations) to other tools; and, how the development of AI in the military domain more broadly will impact the direction and challenges posed by autonomous weapons systems will need consideration.

### **Recommendations**

Technologies incorporating AI and automation into the use of force in armed conflict are currently being deployed without specific agreed rules; the principles various states have proposed and committed to so far have been too weak and vague to prevent civilian harm and risks to peace and security.

All developments in autonomy and AI in the use of force which threaten our safety, security, and humanity must be urgently and adequately addressed through strong regulation by the international community, with unacceptable uses prevented.

States must:

- Move with urgency to negotiate and adopt a new international treaty to prohibit and regulate autonomous weapons systems;
- In International discussions, critically and meaningfully engage with the implications and real-world consequences of current practice in the use of tools that fall under ‘AI in the military domain,’ including acknowledging and examining humanitarian harm;
- Fully consider the legal, ethical, humanitarian, and peace and security risks of further development and use of such systems, whatever the perceived ‘benefits’ may be
- Work urgently to prevent unacceptable uses of technology and trends in development, through committing to develop strong norms for meaningful human control and against digital dehumanisation:
  - This should take place domestically, regionally, and internationally.
  - It must involve a comprehensive and holistic international response, including a legally binding instrument prohibiting and regulating autonomous weapons systems alongside other measures.
  - It should include consideration and development of the other legal instruments necessary to preserve meaningful human control and to protect human dignity against AI in the use of force.

### **Stop Killer Robots Youth Network**

[10 April 2025]

The Stop Killer Robots Youth Network welcomes the opportunity to submit recommendations for consideration by the United Nations Secretary-General in response to Resolution [79/239](#) “Artificial intelligence in the military domain and its implications for international peace and security” adopted by the General Assembly on 24 December 2024. As a global network of young people under age 30 in over 50 countries working to secure a future free of automated killing, we have consistently advocated for the creation of a new treaty on autonomous weapons systems (AWS) – in particular, we insist on a total prohibition of anti-personnel autonomous weapons as we wish to build a world without such dehumanising weapons. While youth will

inevitably face the risks of new weapons technologies, we remain underrepresented in the decision-making process and are often sidelined in forums that shape our interests. As youth who have grown up in an increasingly digital world, we wish to create a future where technology is used to promote peace, justice, equality, and human rights, not perpetuate violence.

With escalating conflicts and the rapid deployment of new weapons technologies around the world, there is an urgent need to reinvest in international law as a measure to build trust and achieve sustainable peace and security. The application of artificial intelligence (AI) in the military domain presents numerous challenges that concern us as youth, including digital dehumanisation, the gamification of violence, and the further erosion of human control and involvement over the use of force.

### **Military AI & AI systems already in use**

Artificial intelligence has been progressively implemented in the military domain over the past decade, however, due to the opacity of military activities and development, the wide public has not been aware of this issue until recently when the active uses of AI systems have been mediatized. We have seen and monitored the use of AI systems to support the targeting of both objects and people. Unfortunately, the use of such systems have not been able to alleviate civilian suffering, for example, in Gaza where one third of victims are children and where too many civilian infrastructures, including critical infrastructures such as humanitarian camps, hospitals<sup>1</sup>, and schools<sup>2</sup>, have been either directly targeted or indirectly impacted by the hostilities.

There have been other concerning uses<sup>3</sup> of AI systems outside of the military which need to be considered as they might be implemented in the military domain, mainly predictive AI and facial recognition. Predictive AI technologies have been used in the police and judicial systems since the early 2010s and have been shown to be ineffective, incorrect, and subject to reinforcing discriminatory behavior.<sup>4</sup> If predictive AI were to be implemented in the military domain, it could lead to the increasing risk of civilians being targeted as they could be labeled as possible fighters or being indirect victims of military activities due to the multiplications of targets with predicted military advantages. Facial recognition technologies (FRTs) are also of concern as they are also unreliable especially when it comes to the identification of non-white males. Facial recognition-enabled targeting in military operations must be prohibited as those systems cannot comprehensively analyse every factor that makes military personnel or civilians a target or not.

### **Digital dehumanisation**

One of the main concerns we have about the use of AI systems in the military domain is the proliferation and banalisation of “**Digital dehumanisation**”. We define digital dehumanisation as the process whereby humans are reduced to data, which is

<sup>1</sup> World Health Organization (2025), ‘oPt Emergency Situation Update’. [https://www.emro.who.int/images/stories/Sitrep\\_57.pdf](https://www.emro.who.int/images/stories/Sitrep_57.pdf).

<sup>2</sup> Save the Children (2025), ‘Education Under Attack In Gaza, With Nearly 90% Of School Buildings Damaged Or Destroyed’. <https://www.savethechildren.net/blog/education-under-attack-gaza-nearly-90-school-buildings-damaged-or-destroyed>.

<sup>3</sup> Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica (2016), ‘Machine Bias’. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

<sup>4</sup> Will Douglas Heaven, MIT Technology Review (2020), ‘Predictive policing algorithms are racist. They need to be dismantled’. <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/>.

then used to make decisions and/or take actions that negatively affect their lives. This process deprives people of dignity, demeans individuals' humanity, and removes or replaces human involvement or responsibility through the use of automated decision-making in technology.<sup>5</sup> Additionally, the increased speed and scale of target production through military AI erodes moral restraints in war and lowers the impact and capacity of decisions from human operators<sup>6</sup>, thus enabling the AI systems to make decisions without meaningful human control, which further dehumanises the decision-making process.

### **Relying on (Big) data leads to problems**

We also believe that the use of (big) data in the military leads to multiple issues which need to be considered.

One of the primary issues is the challenge of data labeling – the process of categorizing and tagging data to train algorithms. Inaccurate or biased labeling can have far-reaching consequences, particularly in the context of distinguishing between combatants and non-combatants in conflict zones. If the data used to train military AI systems is flawed or biased, it can lead to disastrous mistakes, such as the targeting of innocent civilians or misidentification of threats.

A critical issue when relying on big data is that the nature data itself is often broken and is incomplete. This means that the data used to train AI models can be incomplete, outdated, or unrepresentative of real-world situations. Such flaws in data can lead to systems that fail to generalize properly, resulting in inaccurate or incorrect predictions and decisions. For example, in combat situations, a lack of diversity in the data used to identify individuals could lead to inaccurate targeting, with devastating consequences. Important data might be missing or poorly represented, such as the exact location of civilians or combatants, which can lead to AI failing to make informed and balanced decisions. In a war scenario, a system trained with data from a specific past conflict may not be capable of handling a new, unpredictable situation. For instance, an AI system that has been fed data from one particular type of conflict might struggle to apply that data to a war with entirely different characteristics, resulting in errors in target identification or incorrect decision-making.

Another significant problem is that many AI systems operate as black boxes. This means that while these systems make decisions and predictions based on the data they process, the decision-making process is not transparent or easily understood. In military scenarios, where the consequences of decisions are extremely serious, the lack of transparency is particularly concerning. If an AI system makes an error, such as wrongly identifying a civilian as a combatant, the absence of clarity about how the system reached that conclusion makes it nearly impossible to understand the origin of the error. This makes accountability difficult, as we cannot determine why the system acted in a particular way. The lack of explanation regarding the decision-making processes of AI also makes it impossible to correct or adjust the system's behavior, potentially perpetuating errors without the ability to fix them effectively.

---

<sup>5</sup> Automated Decision Research (2022), 'Autonomous weapons and digital dehumanisation'. <https://automatedresearch.org/news/report/autonomous-weapons-and-digital-dehumanisation-a-short-explainer-paper/>.

<sup>6</sup> Marta Bo and Jessica Dorsey, OpinioJuris (2024), 'Symposium on Military AI and the Law of Armed Conflict: The 'Need' for Speed – The Cost of Unregulated AI Decision-Support Systems to Civilians'. <http://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-the-need-for-speed-the-cost-of-unregulated-ai-decision-support-systems-to-civilians/>.

Linguistic and cultural bias embedded in data which is used to train AI systems can create security vulnerabilities and catastrophically misinterpret communications, behaviors, and intentions across diverse cultural contexts, potentially triggering lethal automated responses to misunderstood signals.<sup>7</sup> These systems risk automating and amplifying existing prejudices at unprecedented scale and speed with life or death consequences in conflict zones where cultural misunderstandings could rapidly escalate into devastating military actions causing dire consequences.

### **Accountability**

The inclusion of AI systems in the command and decision-making chains will indubitably lead to a lack of accountability and liability for those relying on these systems to make decisions. It will create a sense of distance and lack of liability on the consequences of a decision which mean that decisions may be made without specific, consistent and thorough analysis of the lawfulness and humane characters of the decision. Then, if an action taken using AI systems violates IHL, the people involved in the implementation and those involved in the decision-making should be held accountable and the use of an AI system shall never exempt people from their responsibilities.

We recognize that military operations are bound by multiple bodies of law – national law, International Humanitarian Law (IHL) and International Human Rights Law (IHRL) – which need to be respected and implemented in order for operations to be lawful. Unfortunately, rules of engagement and of targeting – and all the exceptions – cannot be fully understood and implemented by AI systems. Concepts like doubt, proportionality, and the balance between humanity and necessity are inherently human judgments that cannot be captured by an algorithm. Machines cannot be trusted to uphold these standards on their own. Therefore, it is critical that AI systems never act in a vacuum and that humans retain oversight and decision-making power at all times.

### **What the future might look like**

While AI theoretically has the potential to enhance precision and efficiency in military operations, its integration into warfare raises significant concerns about the future of global security. Autonomous weapons systems, capable of making life-or-death decisions without human control, introduce ethical dilemmas and risks of unintended consequences. The use of AI in military technology is likely to aggravate the existing arms race, as nations compete to develop increasingly sophisticated AI systems, widening the power gap between technologically advanced countries and those less developed, leaving them vulnerable in terms of military readiness. The deployment of autonomous weapon systems and AI-driven tools makes conflict more unpredictable, scalable, and asymmetric, granting certain nations the ability to unleash devastating technologies that smaller states or non-state actors may not be able to counter. The proliferation of AI in the military sphere also raises the threat of terrorism, as organized actors could easily access advanced AI-powered systems. Moreover, the fast-paced, constantly evolving nature of AI development turns military strategies into a “cat and mouse” game, where advancements are met with equally rapid countermeasures. In light of these challenges, the future of military AI must be handled with extreme caution, emphasizing robust ethical frameworks, international regulations, and stringent human oversight to prevent these technologies from destabilizing global peace.

<sup>7</sup> Jimena Sofia Viveros Álvarez, Humanitarian Law & Policy (2024), ‘The risks and inefficacies of AI systems in military targeting support’. <https://blogs.icrc.org/law-and-policy/2024/09/04/the-risks-and-inefficacies-of-ai-systems-in-military-targeting-support/>.



## What we need

We call for the establishment of a meaningful legally binding instrument for the use of AI-driven systems in the military requires comprehensive integration of the technical sector alongside state actors, addressing the urgent need for standardized verification protocols and trust-building mechanisms between nations. Such an instrument should define clear autonomy thresholds that specify permissible levels of independence in target selection and engagement, mandate extensive documentation of algorithmic decision processes and testing methodologies and establish explicit red lines that cannot be crossed including prohibited deployment scenarios, target categories, and operational environments. This framework should apply consistently across developing and developed nations, incorporate independent verification bodies with appropriate technical expertise to conduct regular compliance audits, and establish enforcement mechanisms with meaningful consequences for violations, all while facilitating technical data sharing and research that builds confidence between stakeholders in this domain.

These systems present an unprecedented threat to global security and human rights, and the risks they pose to non-combatants are immense. It is crucial that it implements a robust framework of monitoring, accountability and oversight. Firstly, the states need to be bound by positive obligations to ensure the responsible use of AI in the military domain. Accountability is a fundamental aspect of this framework. We call for comprehensive mechanisms that oversee every stage of the AI system life cycle, from development and updates to transfers and research. States must ensure that any uses of AI systems are monitored, with clear reporting structures in place to address incidents promptly. Furthermore, it is vital that human operators using these systems receive thorough training and guidance to make ethical decisions in the field. The principle of meaningful human control must remain central when it comes to the use of AI in the military domain to ensure that ultimate responsibility for any actions remains with human decision makers.

## Unione degli Scienziati Per Il Disarmo

[6 April 2025]

### Introduction

USPID (*Unione degli Scienziati Per Il Disarmo, Union of Scientists for Disarmament*) is an association of concerned scientists – founded in 1983 and based in Italy – which promotes arms control and disarmament initiatives based on scientific understanding of risks posed by military applications of science and technology. USPID submits to the United Nations Secretary-General its views on “Artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and its implications for international peace and security”, in accordance with the invitation formulated in operative paragraphs 7 and 8 of Resolution [79/239](#) adopted by the UN General Assembly on 24 December 2024.

### Hazards for peace and security arising from AI military applications

USPID expresses its deep concern about new hazards for peace, international security, and the respect of International Humanitarian Law (IHL) which arise on account of the ongoing and accelerating military efforts to incorporate Artificial Intelligence (AI) into multiple facets of warfare. Major sources of these hazards have been identified in current limitations of our capability to understand, predict precisely, and control the behavior of AI systems developed by machine learning methods and their interactions with other human or artificial agents. Initially identified in



connection with the operation of AI-enabled Autonomous Weapons Systems (AWS), these hazards are now spreading to AI systems supporting intelligence collection, the achievement of situational awareness, and human decision-making in warfare.

Exceptionally grave concerns are raised by proposals to integrate AI in Nuclear Command, Control, and Communication (NC3) and in adjacent systems supporting nuclear decisions, and to let AI perform tasks that might directly or indirectly affect nuclear decision-making. A significant case in point is the proposal to use AI technologies in nuclear early warning and decision-support systems, which is being advanced with the expectation that AI accuracy will reduce potential errors, and its processing speed will buy more time for nuclear decision makers. However, on account of the probabilistic nature of AI information processing, one cannot exclude the risk of AI perception leading to false positives of a nuclear attack or producing perniciously unreliable recommendations given the impossibility of ensuring that the underlying models are aligned with human values and the UN overarching goal of preventing and removing threats to peace. If such mistakes occur, no matter how infrequent, large-scale and even existential implications for humanity might ensue. Accordingly, it would be imperative to proceed with time-consuming verifications of AI responses in nuclear early warning. But these verifications would be hindered by the black-box nature of much AI information processing and by the reliance on mostly simulated data, eventually thwarting the expectation of buying more time for human decision makers.

Additional concerns are raised by proposals to exploit the rapid pace at which AI operates to speed up battlefield decision-making and targeting cycles. These proposals are fueled by the goal of gaining military advantage over potential adversaries. However, fighting at machine speed jeopardizes both the effectiveness of human oversight on AI-enabled decision support systems and the fulfilment of ethical and legal roles that are attached to human oversight of warfare action. Indeed, overly tight temporal windows for decision-making hinder effective human control over IHL threats raised by machine suggestions. Human interventions which aim at preventing inadvertent conflict escalations prompted by fighting at machine speed are similarly hampered. In addition to this, excessive speed in human-machine interactions has been identified as a factor inducing automation biases on the battlefield, and potentially skewing human decision-making even in the absence of AI failures.

Further hazards arise in connection with inherent vulnerabilities of AI learning methods and systems. Malicious manipulation of input data might be exploited to induce classification mistakes by AI systems. Moreover, poisoning attacks corrupting learning datasets may impair learning processes and the accuracy of resulting AI systems. These risks are compounded by our current inability to fully align AI systems with human goals and values, potentially causing them to deviate from strategic objectives.

### **Recommended actions**

Mindful of these and other emerging hazards posed by the rapid adoption of AI technologies and systems in the military domain, USPID recommends

- to integrate discussion of AI in NC3 into the Non-Proliferation Treaty framework and in dedicated high-level dialogues and forums such as the Summit on Responsible Artificial Intelligence in the Military Domain (REAIM);
- to develop sustained international dialogue, good practices, and confidence-building measures concerning new and emerging risks for peace and IHL respect raised by AI warfare applications;

- to support a comprehensive and detailed inquiry aimed at identifying actual and potential AI applications in the military domain, jointly with situations of use that pose serious threats to peace, international stability, and the respect of IHL;
- to consider and investigate the need to introduce international regulations or prohibitions for those AI military applications that pose serious threats to peace, international stability, and the respect of IHL.

## **Women's International League for Peace and Freedom**

[24 May 2024]

The Women's International League for Peace and Freedom (WILPF) has opposed war and the development of technologies of violence since its founding in 1915. WILPF has consistently condemned military spending and militarism as detrimental to human life and wellbeing. Our concerns with artificial intelligence (AI) in the military domain and its implications for international peace and security are grounded within our wider opposition to weapons, war, and violence, as well as in our opposition to patriarchal, racist, and colonial power relations that are embedded within AI technology.

While there are many perils of the military use of AI; WILPF's submission is focused on the following issues:

1. The need for human emotion, analysis, and judgement in relation to the use of force;
2. The existence of gender, racial, and other bias in AI technology and the implications for digital dehumanisation;
3. The impacts of military use of AI on privacy and personal data;
4. The environmental harms exacerbated by the military use of AI; and
5. The dangers of war profiteering and arms racing.

Due to the concerns raised in this submission and in other spaces, WILPF opposes the military use of AI. This technology, rather than placing limits on violence or harm, expands both. Governance is insufficient in the face of the profits and power the developers of these technologies seek.

In light of the concerns raised in WILPF's submission and the implications for international peace and security, WILPF urges states:

- To refrain from using AI in the military domain and to develop national laws and regulations to this end;
- To pursue a global prohibition on the military use of AI;
- To not develop autonomous weapon systems or AI-enabled weapon systems, including those that can be used to target human beings;
- To ensure protection of personal data from use by militaries, police, border enforcement, and private companies and contractors collaborating with these institutions;
- To uphold human rights and dignity online and offline; and
- To address the environmental harms generated by data centres, cloud computing, and AI by reducing the number of these centres and energy consumption and water use, which will include reducing the overall use of AI.

WILPF also urges:

- Technology companies, tech workers, scientists, engineers, academics and others involved in developing AI or robotics to pledge to never contribute to the development of AI technologies for military use;
- Financial institutions such as banks and pension funds to pledge not to invest money in the development or manufacture of AI for military use; and
- Activists, academics, affected communities, and other concerned about privacy rights, digital dehumanisation, environmental and climate justice, gender-based violence, and other issues to collaborate and strategise to oppose the development and use of AI in the military and other violent domains.

## D. Scientific Community

### AI, Automated Systems, and Resort-to-Force Decision Making Research Project, The Australian National University

[11 April 2025]

#### Introduction

This executive summary highlights policy recommendations outlined in *AI, Automated Systems, and Resort-to-Force Decision Making – Policy Recommendations: Submission to the UN Secretary General Pertaining to A/RES/79/239 (11 April 2025)*, available on the UNODA website. For a complete account of the underlying research and associated research papers, please refer to the full submission.

#### Underlying Research Project

This research has arisen from a **two-and-half-year research project (2022–2025)**, entitled *Anticipating the Future of War: AI, Automated Systems, and Resort-to-Force Decision Making*, led by Professor Toni Erskine (Australian National University) and funded by the Australian Government through a grant by the Department of Defence.

Its focus is **distinctive and critical**. While the attention of academics and policy makers has been overwhelmingly directed towards the use of AI-enabled systems in the *conduct of war* – including, prominently, on the emerging reality of ‘lethal autonomous weapons systems’ (‘LAWS’), this project has addressed the **relatively neglected prospect of employing AI-enabled tools at various stages and levels of deliberation over the resort to war**. In other words, ‘it takes us from AI on the battlefield to **AI in the war-room**’.<sup>1</sup>

This research project has brought together **leading scholars and practitioners** working on different aspects of international politics and security, strategic and defence studies, and artificial intelligence (AI) to contribute to a multi-disciplinary study and set of **policy recommendations on the risks and opportunities of introducing AI, machine learning (ML), and automated systems** into state-level decision making on the **initiation of war**. Our interventions are made from the perspectives of political science, international relations, law, computer science, philosophy, sociology, psychology, engineering, and mathematics.

Project participants presented and discussed their research at two workshops (June 2023 and July 2024) at the Australian National University (ANU), convened by

<sup>1</sup> T. Erskine and S. E. Miller, ‘AI and the Decision to Go to War: Future Risks and Opportunities’, *Australian Journal of International Affairs*, Vol. 78: 2 (2024), pp. 135–147 (p. 138).

Professor Toni Erskine and Professor Steven E. Miller (Harvard). Participants also received feedback on their initial research-based policy recommendations from senior Australian Government delegates from the federal civil service as part of a one-day policy roundtable (July 2024) at the ANU.

### ‘Four Complications’

For all the potential **benefits** of AI-driven systems – which are able to analyse vast quantities of data, make recommendations and predictions by uncovering patterns in data that human decision makers cannot perceive, and respond to potential attacks with a speed and efficiency that we could not hope to match – challenges abound. Through this project, we have sought to address **four thematic ‘complications’** that we propose will accompany the gradual infiltration of AI-enabled systems in **decisions to wage war**:<sup>2</sup>

- **Complication 1** relates to the displacement of human judgement in AI-driven resort-to-force decision making and possible implications for deterrence theory and the unintended escalation of conflict.
- **Complication 2** highlights detrimental consequences of automation bias, or the tendency to accept without question computer-generated outputs – a tendency that can make human decision makers less likely to use (and maintain) their own expertise and judgement.
- **Complication 3** confronts algorithmic opacity and its potential effects on the democratic and international legitimacy of resort-to-force decisions.
- **Complication 4** addresses the likelihood of AI-enabled systems impacting organisational structures and chains of command, whether degrading or enhancing strategic and operational decision-making processes.

Contributors to this project have explored these proposed complications in the context of either **automated self-defence** or the use of **AI-driven decision-support systems (DDS)** that would inform human resort-to-force deliberations. We have identified risks and opportunities of using AI-enabled systems in these contexts and make recommendations on how risks can be mitigated and opportunities promoted.

### Complication 1: Displacement of human judgement

#### AI in Nuclear Crisis Decision Making

One key area of research undertaken in response to this complication is the nuanced interplay between AI and human decision making in the high-stakes context of **nuclear crisis management**. **Risks (including the increased fragility of nuclear deterrence relationships, crisis signalling becoming more complex, and unintended escalation)** have been explored in two broad areas: i) automation in military deployments, or taking the human ‘out of the loop’ in the decision to use nuclear or strategic non-nuclear weapons (SNNW); and, ii) the integration of AI into human decision-making (particularly in early warning threat assessments). Although much of this research has focused on risks, **novel benefits** of introducing AI-driven decision-support systems (DSS) into human-led nuclear crisis management have also been proposed.

### Policy Recommendations:

---

<sup>2</sup> For an account of these ‘four complications’, see T. Erskine and S. E. Miller, ‘AI and the Decision to Go to War: Future Risks and Opportunities’, *Australian Journal of International Affairs*, Vol. 78: 2 (2024), pp. 135–147 (pp. 139–40).

- **Always incorporate human-in-the-loop safeguards:** Ensure AI systems in nuclear command and control are always overseen by human operators and that human decision-making remains central to determining when and how nuclear-weapon states resort to the use of their arsenals.
- **Promote a holistic approach to AI-safety:** AI safety should account for both technical and socio-technical dimensions. Assess safety challenges in AI-enabled DSS comprehensively, including issues of security, trust, and liability.
- **Broaden the scope of risk assessments:** Apply risk assessments relating to the deployment of AI and ML not only to obvious areas such as nuclear launch orders, but also to less obvious areas such as early warning intelligence assessments (including by non-nuclear allies) and SNNW capabilities (including by non-nuclear allies).
- **Restrict the use of AI-assisted warning data:** The key to balancing the benefits of incorporating AI into early warning against the risks is limiting what AI-assisted warning data is used for. In AI research, prioritise tasks such as calculating effective evasive manoeuvres in the event of an attack and using pattern recognition and anomaly detection to improve arms control verification.
- **Pursue informal arms control and confidence-building:** Advance informal measures such as regular dialogue, red-line agreements, and information-sharing mechanisms. Expand unilateral initiatives like moratoriums where feasible.
- **Explore AI's potential to promote empathy and enhance decision making:** Decision makers must exercise 'security dilemma sensibility' (SDS) in times of crisis. Decision makers and diplomats exercise SDS when they are open to the possibility that the other side is behaving the way they are because they are fearful and insecure, and crucially, recognize the role that their own actions may have played in this. Explore ways that the balanced integration of AI and human judgement could enhance SDS during nuclear crises by promoting empathy and trust.

### **AI Mistakes in the Resort to Force**

Another area addressed in relation to this complication is **state responsibility** when **errors** occur in AI-driven or autonomous systems involved in resort-to-force decisions. Such errors may arise from poor system training, data poisoning by adversaries, or two AI-driven systems interacting in unintended ways. It is essential to develop legal standards and practices that reduce the risk of unintended conflict resulting from such failures.

### **Policy Recommendations:**

- **Adopt robust security and cyber hygiene:** States should adopt robust protections against AI data poisoning and cyber attacks to meet jus ad bellum standards of good faith and reasonable conduct.
- **Clarify legal guidelines on delegating the use of force to autonomous systems:** Senior leadership within states should set clear domestic legal standards regarding when and how autonomous systems may be authorised to use force.
- **Commit to transparency in after-action reviews:** States should commit to being transparent and deliberate about after-action reviews of any AI errors that occur in the field, potentially drawing on civilian casualty review processes as a model.

### Complication 2: Automation bias

Our research in response to the second complication focuses on the relationship between human actors and **AI-driven DSS** in resort-to-force decision making. It includes a detailed survey-based study of **military trust in AI** during strategic-level deliberations and a robust account of the importance of ensuring that there are human **‘experts-in-the-loop’** when AI-driven systems contribute to decisions on war initiation. This body of work also addresses the **benefits** of employing DSS to **enhance our cognitive capacities** in strategic decision making and, conversely, uncovers the potential **dangers** of such reliance if these systems **dull our sensitivity to the tragic qualities of war** or contribute to the **erosion of restraint** by creating the illusion that they replace us as responsible actors.

#### Policy Recommendations:

- **Consider the multidimensionality of trust:** Recognize that soldiers’ trust in AI is not a forgone conclusion. Rather, it is complex and multidimensional, and further complicated by biases, uncertainty, and lack of education.
- **Interrogate norm compliance:** In terms of governance, explain how policies on increasingly autonomous capabilities coincide or diverge from international norms and laws informing their use.
- **Embed experts in decision structures:** Enshrine an ‘expert-in-the-loop’ organisational structure – i.e., high-level experts as core decision makers.
- **Prohibit automation:** Prohibit automation of resort-to-force decisions.
- **Increase AI literacy of domain experts:** Provide and require basic technical training for high-level domain experts so they understand the logics of AI and can thus incorporate AI decision inputs from an informed position.
- **Provide on-going, substantive training for domain experts:** Sustain substantive training for, and assessment of, high-level experts to bolster and ensure substantive competencies.
- **Regulate non-autonomous AI:** While autonomous AI agents, e.g., lethal autonomous weapons systems (LAWS), need regulation, so do non-autonomous AI systems, which leave humans vulnerable to new forms of influence, moral and cognitive atrophy, and undermined responsibility.
- **Design AI-driven DSS to promote more accurate perceptions of their capacities:** Ensure AI-driven DSS are not easily mistaken for responsible agents in themselves by avoiding anthropomorphism, building in warnings about system limitations, and incorporating features that emphasise human agency and accountability.

### Complication 3: Algorithmic opacity

Our research in response to the third complication addresses **how the lack of transparency of AI-driven decision making can threaten the legitimacy** of AI-informed decisions on the resort to force. This body of work includes original research on **large language models (LLMs)** and their potential to exacerbate existing **pathologies in intelligence analyses**. It also examines the role that the **‘architecture of AI’** and its hidden vulnerabilities play in deliberations surrounding the resort to force. Moreover, research within this pillar conceives of military decision-making institutions as **‘complex adaptive systems’** – a conceptual framework that yields a range of insights, including that human-machine teams possess a form of **‘cognitive diversity’ that could be leveraged for more efficient decision-making**, but also

exploited to poison information flows, and that technical explanations for algorithmic opacity will not solve accountability concerns.

#### Policy recommendations:

- **Develop policy to limit epistemic pathologies of LLMs:** States should clearly determine defence and intelligence policy towards either a) procurement of LLMs, b) state development of LLMs, or c) a combination of both. They should use this guidance to develop policy which seeks to limit the epistemic pathologies of LLMs in autonomous decision-making.
- **Commit to sector-wide procurement guidelines and oversight of generative AI tools** used in decision-making chains.
- **Commit to regulating data markets** and access to those markets through alliance relationships.
- **Promote understanding of the tech ecosystem and its fragilities:** Increase understanding of the inherent interdependencies and vulnerabilities in the tech ecosystem, including by creating technology literacy training programs designed specifically for politicians and policy, intelligence, and military leaders.
- **Invest in research** to develop a comprehensive picture of the architecture – physical and digital – that underpins AI, including critical dependencies and vulnerabilities and how access and power are distributed.
- **Invest in research on social media** and its impact on functions of government, including its potential to disrupt democracies, facilitate foreign interference, and influence decision making on the use of force.
- **Recognize AI's current influence:** Significantly increase awareness of government reliance on the architecture of AI, especially for critical government functions, including resort-to-force decision making.
- **Invest in research and development** to maximize the benefits of human-machine cognitive diversity.
- **Implement responsible AI governance programs** that carefully balance accountability with operational efficiency.
- **Perform regular red-team exercises** to ensure that the integration of AI in decision-making institutions does not induce systemic blind spots and vulnerabilities in military decision-making.

#### Complication 4: Impact on organisational structures

Our research regarding the fourth complication explores both the **beneficial and damaging effects that AI-driven systems can have on institutional structures** in the context of resort-to-force decision making. Studies focus on how AI-driven DSS can improve 'adaptive culture' within military organisations, thereby improving wartime decisions, and how the urgent need to **upgrade AI literacy and educate human analysts** should lead us to **reform institutional structures and cultures**. The novel notion of '**proxy responsibility**' is proposed as an institutional response to ensure that responsibility can be meaningfully assigned to humans for resort-to-force decisions that are informed by AI systems. Moreover, original research highlights the significance of the **neglected category of AI 'integrators'** – sandwiched between the 'developers' and 'users' of AI within organisational structures – when it comes to strategic military applications of AI.

**Policy recommendations**

- **Set (and evolve) measures of effectiveness.** If AI-enabled adaptive capacity is to work effectively, measures of military effectiveness must guide which direction adaptation might take. Establish such measures at the tactical (battlefield) and strategic (war-room) levels to guide development and implementation of AI-enabled adaptation.
- **Know where adaptation relevant data is found, stored and shared.** An enhanced adaptive stance in military institutions must have enhanced data awareness as a foundation. Data awareness and management must become one of the basic disciplines taught to military personnel.
- **Scale AI support from individual to institution.** There is unlikely to be a one-size-fits-all algorithm or process that can enhance learning and adaptation at every level of military endeavours. Create a virtual 'arms room' of adaptation support algorithms as part of an institution-wide approach to adaptation.
- **Routinely question AI-enabled outputs:** Build mindsets, protocols, institutional cultures, and inter-agency structures in 'normal' pre-crisis times to routinely question AI-enabled output from human-machine teams.
- **Institute an advisory body:** In order to support the notion of 'proxy responsibility' as an institutional response to 'responsibility gaps' when decisions on war initiation are informed by AI-enabled systems, establish and/or strengthen state-level 'AI departments'. These departments would integrate technical, political, and ethical competence and expertise and advise on resort-to-force decision-making processes.
- **Support research on AI integration:** Fund research on the integration of AI in strategic decision-making.
- **Provide standards:** Outline minimum standards for the responsibilities of AI developers and integrators.
- **Facilitate inter-group discussions** between developers, integrators and users during development, integration, and longer-term maintenance processes.
- **Create accountability guidelines:** Provide well-defined guidelines and rules indicating who is accountable if something goes 'wrong'.



---

**Queen Mary University of London, T.M.C Asser Institute,  
University of Southern Denmark, University of Utrecht**

[11 April 2025]

Views of members of the scientific community and civil society; specifically, we are a group of academics with expertise in ethical, legal and political dimensions of military Artificial Intelligence and herewith put forward our shared views pursuant to resolution 79/239 “Artificial intelligence in the military domain and its implications for international peace and security” adopted by the General Assembly on 24 December 2024, in accordance with the request of the UN Secretary-General contained in Note Verbale ODA/2025-00029/AIMD.

**Introduction:**

The rapid advancement and integration of AI technologies into targeting operations have sparked ongoing debates surrounding their ethical, legal, and operational implications. Over the past decade, the discourse on AI in warfare has largely centered on autonomous weapon systems (AWS),<sup>66</sup> driven in part by the initiation of discussions in 2013 and the formalization of a regulatory process under the UN Convention on Certain Conventional Weapons (CCW) and the Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE LAWS), which exclusively focuses on lethal AWS.<sup>67</sup> However, the increasing integration of AI-based decision-support systems (AI-DSS) into targeting practices<sup>68</sup> introduces new layers of complexity that demand closer attention from a broad range of stakeholders. This submission responds to that need, structured around three key components: (1) a brief overview of how AI-DSS are currently used in targeting decisions; (2) an analysis of key concerns, including how these systems shape the potential exercise of human judgement and control and underline fundamental gaps in global governance; and (3) a concluding set of recommendations.

**1. Overview of AI-DSS and the joint targeting cycle**

Defined as “the process of selecting and prioritizing targets and matching the appropriate response to them, considering operational requirements and capabilities,”<sup>69</sup> targeting is a core military function at the very heart of warfare. While the potential range of use cases for AI-DSS in military decision-making is broad, in

---

<sup>66</sup> The latest definition of AWS from the CCW GGE LAWS Rolling Text (26 November 2024): “A lethal autonomous weapon system can be characterized as an integrated combination of one or more weapons and technological components that enable the system to identify and/or select, and engage a target, without intervention by a human user in the execution of these tasks.” On file with authors.

<sup>67</sup> For a brief overview of some of the latest developments of the GGE LAWS see Jeroen van den Boogaard, *Warning! Obstacles Ahead! The Regulation of Autonomous Weapons Systems in the GGE LAWS*, *Opinio Juris*, 4 March 2024 found at: <https://opiniojuris.org/2024/03/04/warning-obstacles-ahead-the-regulation-of-autonomous-weapons-systems-in-the-gge-laws/>.

<sup>68</sup> There have been several reported uses of AI-DSS by Israel in Gaza and potentially in Lebanon, by both Ukraine and Russia in the ongoing conflict, and by the United States in its actions against Houthi rebels in the Red Sea and in Yemen, to name a few. For a comprehensive overview of literature in this space, see e.g., Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates*, Centre for War Studies, University of Southern Denmark, November 2024. Found here: <https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/>.

<sup>69</sup> United States Department of Defense, *Dictionary of Military and Associated Terms*, March 2017, found at: <https://www.tradoc.army.mil/wp-content/uploads/2020/10/AD1029823-DOD-Dictionary-of-Military-and-Associated-Terms-2017.pdf>.

---

targeting, AI-DSS can be understood to serve as **tools** that use AI techniques to collect and analyze data, provide information about the operational environment as well as actionable recommendations, with the aim of aiding military decision makers in evaluating factors relevant to legal compliance such as taking precautions and ensuring proportionality in attacks.

More specifically, AI-DSS are increasingly integrated across multiple phases of the joint targeting cycle (JTC), including within target development and prioritization, capabilities analysis, and mission execution. The JTC is a reflective example of a structured process used by military forces to identify, evaluate, and engage targets while ensuring compliance with operational, legal, and ethical standards,<sup>70</sup> generally consisting of six (non-linear) phases:

1. **End-State and Commander's Objectives:** Defining strategic military goals and desired outcomes.
2. **Target Development and Prioritization:** Identifying, verifying/validating, and prioritizing targets based on intelligence and mission goals.
3. **Capabilities Analysis:** Assessing the available strike options and their effectiveness.
4. **Force Assignment:** Allocating specific military assets (e.g., airstrikes, artillery, cyber operations) to engage the target.
5. **Mission Execution:** Carrying out the targeting operation while ensuring compliance with relevant laws and the rules of engagement.
6. **Assessment:** Evaluating the effectiveness of the operation and adjusting for future operations, if necessary.

Within this framework, AI DSS are assumed to serve primarily as informational and analytical tools which support human decision-making rather than supplant it. However, this assumption and framing obscures how AI-DSS influence human cognitive processes within the JTC. This impact on human decision-making is often underestimated and remains insufficiently examined, leaving critical discussions about the role of AI-DSS largely absent from current policy debates.

## 2. Analysis of Key Concerns

### (a) (Meaningful) Human Judgement and Control

AI-DSS are often portrayed as enhancing human decision-making and the quality of decisions therein. The perception of AI-DSS as mere subsidiary tools has led to a narrative that the integration of AI-DSS poses fewer challenges than AWS, given that these systems do not directly “engage” targets (i.e., they do not have an inherent capability to directly carry out the use of force) and are tools that assist human commanders. The outputs are ostensibly ultimately reviewed through (several layers of) human oversight, such as processes of verifying and validating targets using

---

<sup>70</sup> Michael Schmitt et al, *Joint and Combined Targeting: Structure and Process*, Chapter 13 in Jens David Ohlin (ed) *Weighing Lives in War* (Oxford, 2017). See also, Jessica Dorsey and Marta Bo, *AI-Enabled Decision-Support Systems in the Joint Targeting Cycle: Legal Challenges, Risks, and the Human(e) Dimension*, forthcoming 2025, *International Law Studies*. “Targeting generally involves four key steps: (1) objectives and guidance, (2) planning, (3) execution, and (4) assessment. Encapsulating these four key steps, the United States and NATO outline their targeting processes through similar six-phase cycles [addressed in this submission]. As the reader can discern, different states employ different doctrines for targeting. What is important ... is not necessarily the specific labels for various steps followed by any given state, but rather how and when compliance with the principle[s of IHL are] incorporated into the targeting process.”

additional intelligence sources.<sup>71</sup> As a result, errors or inaccuracies in AI-DSS outputs are often seen as non-critical, based on the assumption that robust human oversight and appropriate control will compensate for them. However, closer examination reveals that this control is frequently superficial, offering only the appearance of, rather than actual meaningful, or context-appropriate, human judgement and control.

This is because AI-DSS structure and condition the quality of human control and oversight and limit the ways control and oversight can be exercised. The use of AI-DSS creates a shared decision-making space between human military personnel and AI technologies. States appear to have recognized and focused on many of the advantages of this shared decision-making space for military personnel, i.e., how the use of AI-DSS advances human decision-making through offering data-driven insights. But using AI-DSS also delimits the capacity to exercise human oversight and control because of the technologies' complexity and the increased speed (and therefore scale) it can bring to decision-making processes. Rather than supporting human oversight, using AI-DSS may risk humans becoming little more than reactive cogs in socio-technical systems.<sup>72</sup> Moreover, this configuration risks amplifying adverse human biases, such as automation bias, anchoring bias, or cognitive action bias, to the detriment of exercising qualitatively high levels of human control.<sup>73</sup> Considering AI-DSS as a distinct form of technology therefore reveals significant challenges associated with military AI and human oversight, challenges that extend beyond those that arise when simply integrating the technology in weapon systems.

Recent conflicts have shown the risks associated with AI-DSS being employed in critical functions, such as target selection and even nomination, and their conditioning and constraining of human involvement, affecting the fulfilment of core legal obligations embedded within the JTC. The use of AI-DSS raises fundamental concerns about whether human decision makers can retain adequate cognitive autonomy over the JTC process or whether humans will become overly reliant on algorithmic outputs for critical judgements in the context of armed conflict.<sup>74</sup> Consequently, there are significant legal concerns regarding the effects of such systems on decision-making processes and use of force decisions and ability for users to comply with IHL obligations, especially with respect to the obligation to take all feasible precautions to minimize civilian harm to the greatest extent possible in attack and comply with the principles of distinction and proportionality.<sup>75</sup>

Importantly, these concerns are not new. There is extensive debate around how to preserve meaningful human judgment and human agency when conducting IHL-evaluative legal assessments, in the context of AWS. These discussions – which include expert analysis on accountability, human-machine interaction, automation

---

<sup>71</sup> Alexander Blanchard and Laura Bruun, *Automating Military Targeting: A Comparison Between Autonomous Weapon Systems and AI-Enabled Decision Support Systems*, Stockholm International Peace Research Institution (SIPRI) forthcoming 2025 (draft on file with authors).

<sup>72</sup> Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025).

<sup>73</sup> Dorsey and Bo, *supra* n. 5.

<sup>74</sup> *Ibid*; see also Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates*, Centre for War Studies, University of Southern Denmark, November 2024. Found at: <https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/>.

<sup>75</sup> Article 57 of the First Additional Protocol to the Geneva Conventions. See also Dorsey, Bo *supra* n. 5 (on AI-DSS and their effects on the principle of precautions); Jessica Dorsey, *Proportionality under Pressure: The Effects of AI-Enabled Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*, forthcoming *International Review of the Red Cross* (2025) (on AI-DSS and their effects in the context of IHL proportionality assessments).

---

bias, and the effect of AI systems on legal and ethical reasoning<sup>76</sup> – provide valuable lessons that must inform discussions around military AI and specifically the use of AI-DSS in military contexts.

**(b) AI-DSS: Understudied, Under-Addressed and Unregulated**

Framing AI-DSS as mere tools, has led to an underestimation and lack of analysis on the way their use affects the cognitive decision-making process within the JTC. The relative lack of attention paid to AI-DSS so far can partly be attributed to the fact that such systems are seen to be used with a human *in* or *on* the loop framework, with their outputs ostensibly reviewed by one or more individuals during the targeting process. As a result, current understandings of AI-DSS use appear to align with widely supported principles of human control and oversight. However, this gap in the debate is also caused by a lack of transparency around how specific AI-DSS function, and a consistent failure to comprehensively examine how they are being used in practice.

Additionally, the persistent focus on AWS at the expense of AI-DSS obscures the growing reliance on AI in shaping operational and strategic outcomes. Unlike AWS, which have been debated in the framework of the CCW for the past decade, AI-DSS lack a comparable institutional platform. Attention to AI-DSS remains scattered across various initiatives but these efforts have yet to provide the dedicated regulatory focus or coordination needed.

**3. Recommendations:**

- i. **Reassert** the central role of human cognitive and legal reasoning in military operations by implementing safeguards that ensure key legal assessments remain grounded in human(e) judgment. Leverage existing insights from debates on AWS and research on human-machine teaming and human-computer interaction to inform discussions on AI-DSS.
- ii. **Recognize** and address the incremental effects of AI-DSS design and use on human cognitive reasoning and critical deliberation. Promote awareness and attentiveness as a crucial part of reasserting and strengthening the exercise of human agency in targeting decision-making.
- iii. **Reinforce** calls for greater attention to the implications of AI-DSS in armed conflict. Utilize platforms such as the UN General Assembly's First Committee on Disarmament and International Security and the Global Commission on the Responsible Use of AI in the Military Domain to foster inclusive and complementary discussions on the associated risks and systemic changes AI-DSS introduce.

---

<sup>76</sup> Marta Bo, *Autonomous Weapons and the Responsibility Gap in light of the Mens Rea of the War Crime of Attacking Civilians in the ICC Statute*, 19 Journal of International Criminal Justice 2021; Bo, M., Bruun, L. and Boulanin, V., *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS* (SIPRI: Stockholm, Oct. 2022), p. 41; Boulanin, V., Bruun, L. and Goussac, N., *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction* (SIPRI: Stockholm, June 2021), p. 54; and Bruun, L., Bo, M. and Goussac, N., *Compliance with International Humanitarian Law in the Development and Use of Autonomous Weapon Systems: What Does IHL Permit, Prohibit and Require?* (SIPRI: Stockholm, Mar. 2023), p. 24. Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapons systems," *Humanitarian Law and Policy*, 29 August 2018, found at: <https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/>.

---

## United Nations Institute for Disarmament Research

[11 April 2025]

Artificial intelligence (AI) is rapidly transforming the military domain and profoundly influencing international peace and security. Initiatives such as the summits on Responsible AI in the Military Domain (REAIM) and the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, while not being universal processes, have significantly elevated international attention on the military applications of AI. In particular, they have moved the debate beyond lethal autonomous weapon systems (LAWS) and have successfully highlighted the multifaceted impacts of AI, fostering broader international policy engagement. Building on the political momentum generated by these initiatives, resolution [79/239](#) adopted by the United Nations General Assembly in December 2024 represented a significant milestone as the first UN resolution on AI in the military context and has offered Member States, international and regional organizations and the multi-stakeholder community the opportunity to share their views on opportunities and risks.

For many years, the United Nations Institute for Disarmament Research (UNIDIR) has played an important role in shaping and informing discussions on the broader impact of AI in the military domain, both within and beyond applications of this technology in weapon systems. It has undertaken research, facilitated multilateral dialogues, and offered policy insights that underline AI's transformative potential for international peace and security. This policy note draws from all the work conducted to summarise opportunities and risks and to offer a potential roadmap for future policy action.

The international community can now shape how AI is used in the military domain, putting principles of responsible AI at the core. A central challenge is the complexity of defining the “military domain”. States and regions interpret the scope of this domain differently based on their unique security landscapes, realities and operational practices. For some countries, military roles extend to internal security tasks such as policing, border control, combating organized crime, protection of critical infrastructure or humanitarian relief in response to natural disasters. Others maintain a stricter definition, limiting military functions to battlefield engagements. These variations, rather than serving as barriers, offer important context for multilateral discussions. International governance frameworks should remain flexible and inclusive, acknowledging and adapting to diverse national and regional security perspectives.

In the many operational contexts within the military domain, AI acts as a force multiplier across several military tasks, including command and control (C2), information management and intelligence, advanced autonomy, logistics, training and simulation, and organizational and support functions. In C2, AI enhances the speed and quality of decision-making, thereby helping commanders rapidly analyse battlefield scenarios. It has the potential to improve adherence to international humanitarian law (IHL), for example by integrating detailed proportionality and other legal assessments. AI-driven intelligence tools analyse large volumes of data at speed, and so improve situational awareness and threat detection. In logistics, AI optimizes supply chains and predictive maintenance, enhancing operational readiness and improving the sustainability of military operations over time. AI further supports advanced autonomy in drones, cybersecurity, and operations in the information domain. Training and simulation benefit from AI by creating personalized, realistic synthetic environments and scenarios. In short, if developed, deployed and used responsibly, AI could increase operational effectiveness while offering new ways to mitigate risks and reduce harm.

However, integrating AI in military contexts also presents significant risks and challenges – technological, security, legal, policy and ethical.

Technologically, military AI systems face issues related to the quality, availability and inherent biases of data. These may lead to unpredictable and

---

potentially harmful outcomes, including violations of international law. The “black box” nature of AI systems, often coupled with their adaptiveness and highly context-dependent nature, complicates trustworthiness assessments and may, at times, challenge the conduct of effective investigations into alleged violations of IHL. Cybersecurity vulnerabilities also expose AI systems to adversarial attacks, requiring stringent security measures.

Security challenges include risks of miscalculation and unintended escalation, particularly through AI-enabled rapid decision-making processes and AI-enabled autonomy, which may result in escalatory responses. The potential for an AI arms race exacerbates international and regional tensions, possibly leading to destabilizing outcomes similar to historical arms competitions. The proliferation of AI technologies to non-state actors further complicates threat landscapes and necessitates robust life cycle management of military AI systems. Additionally, AI-generated disinformation threatens societal stability by undermining trust in information and can have a direct impact on military operations.

Legal challenges revolve around ensuring compliance with international law, particularly IHL and international human rights law. Key debates focus on, among other things, accountability and both state and individual responsibility for AI-driven actions, especially regarding lethal decisions. States diverge on whether existing legal frameworks are sufficient or if new, specialized regulations are required. Beyond international law, ethical considerations emphasize maintaining human judgment in critical decision-making and preventing societal biases from infiltrating AI systems. The latter requirement calls for greater diversity and inclusivity in AI development. Additionally, bridging gaps between government, academia and the private sector remains challenging yet crucial for effective governance.

Addressing these challenges requires a comprehensive road map with actions at the multilateral, regional and national levels.

Multilaterally, establishing a United Nations-led comprehensive platform that enables a regular institutional dialogue to address military AI’s broader implications on international peace and security is key as it would provide an institutional framework to advance policy discussions. This platform could build on the existing internationally developed AI principles and frameworks, such as UNESCO’s recommendations or the commitments made in the Global Digital Compact (e.g. safe, secure and trustworthy AI) and further refine them for application in the military domain. These principles could be further developed into voluntary norms of responsible behaviour in the development, deployment and use of AI in the military domain and provide a solid foundation for future multilateral instruments. In addition, such platform could be leveraged to develop practical confidence-building measures (CBMs), lead inclusive multi-stakeholder engagement, and deliver global capacity-building programmes that enhance global security via transparency, cooperation and predictability.

Regionally, existing organizational frameworks can be used to tailor CBMs and guidelines that reflect local security contexts. Cross-regional dialogues would facilitate mutual learning, prevent information silos, and include diverse perspective which would encourage globally coherent responses.

Nationally, states should develop comprehensive AI strategies that detail vision, priorities and governance frameworks, ensuring compliance with international norms and ethical standards. Robust governance structures (e.g., dedicated AI steering committees and ethics boards), alongside iterative legal reviews, would enhance accountability and safety. Transparent communication and clearly defined accountability protocols would further support responsible AI implementation. High standards of data governance, life cycle management approaches, rigorous training programmes and updated military operational guidelines complete these proposed national measures, ensuring the responsible integration of AI in the military domain.

Table A below provides an overview of the proposed roadmap for policy action.

**Table A: A roadmap for future policy action**

<i>Level</i>	<i>Action</i>	<i>Rationale</i>
<b>Multilateral</b>	<p><b>Establish a multilateral process under United Nations auspices to provide a comprehensive platform for discussion on military applications of AI and their impact on international peace and security.</b> This process could be leveraged to:</p> <ul style="list-style-type: none"> <li>a. Develop a set of overarching, core principles of responsible AI in the military domain to help align national efforts and reduce risk.</li> <li>b. In the future, further develop these core principles into international voluntary norms or guidelines for responsible state behaviour in the development, deployment and use of AI in the military domain. These guidelines could take the form of a code of conduct or a political declaration supplemented by more technical instruments as required (e.g., on AI assurances, and robust protocols for testing and evaluation).</li> <li>c. Develop confidence-building measures (CBMs) for military AI. States could agree on and implement practical CBMs to increase transparency and trust regarding AI in the military domain.</li> <li>d. Promote multi-stakeholder engagement in support of multilateral policy action.</li> <li>e. Develop and implement a coherent capacity-building programme.</li> </ul>	<p>Collectively, these multilateral actions aim to foster cooperation, set common rules and share knowledge on military AI at the international level with a view to increasing predictability.</p> <p>They aim to shape the global landscape so that all states move towards safer and more transparent integration of AI in the military domain, thereby reducing the risks.</p> <p>While clustered under a single umbrella recommendation, each of the actions above could be implemented on its own, although their mutually reinforcing nature would amplify the impact achieved if they are implemented in combination.</p>
<b>Regional</b>	<p><b>Leverage regional and subregional organizations and dialogues to discuss the issue of AI in the military domain.</b></p>	<p>Regional and subregional approaches allow tailoring to specific security realities and threat perceptions, which could</p>

<i>Level</i>	<i>Action</i>	<i>Rationale</i>
	Regional and sub-regional organizations could:	lead to concrete results that are more aligned with specific needs.
	<ul style="list-style-type: none"> <li>a. Develop region-specific CBMs, norms or guidelines that reflect local contexts.</li> <li>b. Set up networks for information-sharing on AI-related best practices suited to their security landscape.</li> <li>c. Develop joint AI-development projects, aligning operational, legal and technical requirements.</li> </ul>	<p>In addition, regional and subregional approaches could be leveraged to inform and shape global dialogues and strengthen context-specific capacity-building.</p> <p>Cross-regional dialogue can be a useful tool to enable mutual learning and avoid echo chambers.</p>
<b>National</b>	<b>Initiate cross-regional dialogues</b> <b>Initiate cross-regional dialogues on AI</b> , where two or more regional groups exchange lessons and possibly align their approaches.	
	<p><b>Implement a comprehensive approach to AI governance in the military domain</b> to include the following actions:</p> <ul style="list-style-type: none"> <li>a. Develop a comprehensive national strategy or policy on AI in security and defence.</li> <li>b. Establish robust governance structures and review processes.</li> <li>c. Implement transparency and accountability measures</li> <li>d. Implement robust data practices and governance frameworks for all military AI applications.</li> <li>e. Manage AI capabilities throughout their entire life cycle – from design and development, through testing and deployment, to updates and decommissioning – with continuous risk assessments and mitigation at each stage.</li> <li>f. Invest in human capital and training by developing extensive training programmes for military personnel on AI and cultivating a</li> </ul>	<p>A national strategy clarifies roles and responsibilities, and provides a clear direction for the development, acquisition, integration and use of AI in the military domain.</p> <p>Dedicated structures provide focus and accountability. They create effective checkpoints that AI projects must pass and comply with consistently (e.g., ethical approval, legal clearance, safety certification), reducing chances of unsafe or unlawful deployment.</p> <p>Transparency builds public trust and international confidence that a state is using AI responsibly.</p> <p>Accountability ensures that the presence of AI does not create a vacuum of responsibility – maintaining the ethical and legal norm that humans are accountable for military actions.</p> <p>By prioritizing robust data governance and the provision of the necessary infrastructure to enable it, militaries can improve the performance and</p>



<i>Level</i>	<i>Action</i>	<i>Rationale</i>
	new generation of AI-literate officers and specialists. This includes not only technical training but also training on the ethical and legal aspects of AI use in operations.	trustworthiness of their AI systems and reduce error rates.
	g. Review military operational guidelines to strengthen AI governance in military contexts, including military documentation (e.g. doctrines, standard operating procedures and others), and rules of engagement.	A life cycle view ensures that safety and compliance are ongoing commitments reducing chances of failure in the field and ensuring that accountability is maintained throughout the system's use.  Human expertise and judgment remain critical. Training reduces misuse and enables more effective human-machine teaming.  Existing military governance tools and instruments can be used to strengthen the governance of AI in the military domain at a more practical, tactical level, thereby offering an impactful complement to the highest levels of governance and the associated obligations emanating from international, regional and national laws and regulations.

## **E. Industry**

### **Microsoft**

[24 May 2024]

Microsoft welcomes the opportunity provided by the United Nations General Assembly resolution [A/RES/79/239](#) on “Artificial Intelligence in the Military Domain and its Implications for International Peace and Security”, and UNODA’s invitation to share perspectives on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain, with specific focus on areas other than lethal autonomous weapons systems.

Our perspectives reflect Microsoft’s deep commitment to our Responsible AI Principles and our Secure Future Initiative, emphasizing cybersecurity, safeguarding international norms, and promoting trust in technology, and our active participation in multi-stakeholder initiatives including the UNIDIR-led Roundtable for AI, Security, and Ethics (RAISE).

#### **I. Opportunities**

Microsoft recognizes substantial opportunities in responsibly applied AI within the military domain, particularly:

- 
- *Enhancing cybersecurity and defense capabilities:* AI significantly strengthens cybersecurity defenses by automating threat detection, enabling faster and more accurate responses to cyber threats. Technologies such as Microsoft Security Copilot illustrate the transformative potential of AI in defense, empowering cybersecurity professionals to identify and mitigate risks efficiently. Initiatives like Microsoft's Zero Day Quest and collaboration with MITRE ATT&CK demonstrate proactive industry efforts to enhance global cybersecurity preparedness and resilience.
  - *Broad spectrum of military applications:* Beyond cybersecurity, responsibly designed AI can significantly enhance efficiency and effectiveness across logistics, command and control systems, intelligence processing, military training, peacekeeping, humanitarian assistance, and disaster relief operations. Diverse applications underscore AI's transformative potential beyond combat scenarios alone.
  - *Improving compliance with international humanitarian law:* AI technologies should improve the accuracy and effectiveness of targeting processes, aiding militaries to better adhere to principles of distinction, proportionality, and necessity. AI should significantly enhance protections for civilians and civilian infrastructure, thereby reducing unintended collateral damage in conflict.
  - *Capacity building and international cooperation:* The adoption of AI in the military domain presents opportunities for global knowledge-sharing and capacity-building initiatives. International partnerships should support developing nations by sharing security capabilities, knowledge, and best practices, thus bridging technological divides and fostering global stability.

## II. Challenges

Microsoft equally acknowledges significant challenges and risks associated with AI applications in the military domain:

- *AI-enhanced cyber threats:* AI has escalated cyber threat capabilities, empowering state-sponsored and criminal actors to carry out increasingly sophisticated cyber operations. These AI-driven threats include advanced phishing campaigns, automated exploitation of vulnerabilities, and adaptive malware, significantly increasing global cybersecurity risks.
- *Risks of escalation and miscalculation:* Integrating AI into military decision-making risks unintended escalation and/or miscalculation. Rapid, automated decision-making processes may inadvertently lower conflict thresholds, amplifying risks of destabilization or accidental conflict.
- *Proliferation and uncontrolled diffusion:* Uncontrolled diffusion, especially through open-source models and decentralized development, heightens the risk of malicious use by both state and non-state actors, including terrorist groups and cyber mercenaries. Increasingly accessible dual-use and proprietary AI systems enable actors even with limited resources can gain access to capabilities that previously required significant investment or expertise, posing additional threats to international security and stability.
- *Algorithmic bias and ethical implications:* Algorithmic biases embedded within AI systems pose ethical and humanitarian concerns. Biases related to gender, race, age, or socioeconomic factors in AI datasets can intentionally and unintentionally perpetuate inequality and discrimination, particularly within sensitive military and security applications.

- 
- *Digital divides and inequality:* Without deliberate policy actions, disparities between developed and developing nations in AI capabilities could deepen, increasing geopolitical tensions and socio-economic inequalities, thus undermining long-term global stability.

### **III. Relevant normative proposals**

Microsoft recognizes several existing and emerging normative frameworks relevant to AI governance in the military domain, including:

- UNIDIR's RAISE initiative, facilitating international multi-stakeholder dialogues and governance proposals.
- The Responsible AI in the Military Domain (REAIM) Summits, emphasizing transparency, accountability, and human oversight at the international level.
- The US Department of Defense Responsible AI Strategy, highlighting responsibility, equitability, traceability, reliability, and governability.
- NATO's Principles of Responsible Use for AI in Defence, emphasizing reliability, governability, and traceability among member nations.

### **IV. Microsoft recommendations**

To maximize opportunities and mitigate the challenges, Microsoft proposes several key recommendations:

- *Establish clear international norms and standards:* Develop explicit international norms and industry standards governing responsible use and development of military AI. These norms should delineate acceptable and unacceptable behaviors, providing robust frameworks to deter misuse and foster transparency and accountability, supported where appropriate by monitoring or compliance mechanisms. AI governance frameworks should explicitly differentiate operational contexts, such as peacekeeping, humanitarian assistance, crisis management, and conflict scenarios, to appropriately address varied ethical, legal, and humanitarian considerations. To ensure continued relevance, such norms should be periodically reviewed and updated to reflect evolving technological developments and operational realities.
- *Ensure human-centric oversight and accountability:* Adopt policies ensuring meaningful human judgment, oversight, and accountability remain central to military decisions involving AI, particularly regarding the use of force. Clear oversight mechanisms and enforceable accountability structures, including rigorous human control and review processes, are necessary to maintain ethical standards, avoid automation bias, and mitigate unintended consequences.
- *Advance secure and transparent AI development practices:* Promote rigorous technical standards and comprehensive life cycle management protocols covering pre-design, development, testing, deployment, operation, acquisition, and decommissioning. Robust vulnerability management, security audits, and transparent development and deployment processes should be integral components, alongside clear capacity-building measures, ensuring AI systems remain secure, responsible, and resilient throughout their operational life cycle.
- *Enhance responsible data governance practices:* Establish clear international guidelines on responsible data governance specifically tailored to military AI applications. Transparent and accountable data management practices addressing collection, sharing, storage, training, and operational usage are

---

crucial for managing dual-use risks, preventing misuse, and maintaining strict compliance with international legal and ethical frameworks.

- *Address and reduce algorithmic bias:* Prioritize addressing algorithmic bias through rigorous testing, transparent data practices, and inclusive AI development processes. Developers and users should establish clear policies to proactively identify, mitigate, and remediate biases, especially when AI systems are deployed in sensitive military or security contexts.
- *Promote responsible innovation and risk-based regulation:* Support regulatory frameworks that are risk-based, outcome-oriented, and balanced, ensuring they encourage innovation while adequately addressing security and ethical risks associated with AI deployment. Industry should advocate for flexible, adaptive regulations that keep pace with technological change, without imposing overly prescriptive or impractical requirements. Industry-led initiatives, such as voluntary codes of conduct, vulnerability disclosure standards, and collaborative red-teaming exercises, should be actively supported and integrated into broader international normative frameworks.
- *Strengthen international governance and alignment:* Support and actively engage in international initiatives, including REAIM Summits and dialogues at the UN General Assembly and UN Security Council. Robust international governance frameworks, characterized by transparency, clear accountability measures, and trust-building mechanisms, are essential for coherent and inclusive approaches to AI governance. Member States and stakeholders should coordinate closely through these forums to reduce fragmentation and ensure global alignment.
- *Support knowledge-sharing and awareness-raising with the UN system:* Encourage and actively contribute to efforts by the UN Secretariat and relevant UN entities to convene meaningful multi-stakeholder expert dialogues, workshops, and knowledge-sharing on AI in the military domain. Exchanges through voluntary contributions, technical expertise, and collaborative initiatives should aim at enhancing global understanding of AI's implications for international peace and security.
- *Strengthen international cooperation and information sharing:* Encourage robust international cooperation, emphasizing real-time threat intelligence sharing and joint attribution mechanisms. Industry actors should actively participate in collective cybersecurity efforts, enhancing global cybersecurity preparedness and response.
- *Foster multi-stakeholder dialogue and collaboration:* Actively participate in and support forums such as RAISE, involving states, international organizations, academia, civil society, and industry. Such inclusive dialogues are essential for mutual understanding, shaping responsible AI practices, and developing collaborative governance structures.

## V. Conclusion

Microsoft is deeply committed to proactive collaboration with Member States, the UN system, industry, and civil society to implement these recommendations swiftly and effectively. Through sustained collective efforts and ongoing engagement in multi-stakeholder initiatives, Microsoft will continue supporting responsible AI governance, innovation, and practices that meaningfully contribute to international peace and security.