Группа правительственных экспертов Высоких Договаривающихся Сторон Конвенции о запрещении или ограничении применения конкретных видов обычного оружия, которые могут считаться наносящими чрезмерные повреждения или имеющими неизбирательное действие

Distr.: General 4 September 2017

Russian

Original: English

Женева, 13–17 ноября 2017 года
Пункт 6 предварительной повестки дня
Рассмотрение различных аспектов новых
технологий в сфере создания автономных
систем оружия летального действия
в контексте целей и задач Конвенции

# Документ «Пища для размышлений»

# Представлено Председателем

### Введение

- 1. Группа правительственных экспертов (ГПЭ) уполномочена Высокими Договаривающимися Сторонами Конвенции о конкретных видах обычного оружия (КНО) на рассмотрение вопросов, касающихся новых технологий в сфере создания автономных систем оружия летального действия (АСОЛД) в контексте целей и задач Конвенции. ГПЭ КНО является сегодня единственным межправительственным форумом в системе Организации Объединенных Наций, который официально занимается этим предметом, объединяя соответствующие междисциплинарные знания и опыт. Для того чтобы процесс КНО имел ценность и сохранял актуальность, важно иметь уверенность в неизменной финансовой поддержке совещаний ГПЭ и добиваться того, чтобы все Высокие Договаривающиеся Стороны и другие заинтересованные субъекты, включая промышленность, вносили свой вклад в дискуссии.
- 2. Учитывая то, что объем времени, выделяемый для ГПЭ на 2017 год, был сокращен с 10 до 5 дней, Председатель считает, что ГПЭ может хорошо начать работу в этом году с проведения обстоятельного обзора текущего состояния разработок в сфере технологий и их практического применения в конкретных военных системах. ГПЭ могла бы также продолжить проводившиеся в последние три года полезные дискуссии по правовым и этическим вопросам, имеющим отношение к АСОЛД, включая новые изменения в национальном законодательстве, а также в корпоративной сфере, призванные регулировать использование гражданских автономных систем, таких как самоуправляющиеся транспортные средства. Обстоятельный трехкомпонентный обзор технологий, военного воздействия и правовых/этических аспектов помог бы ГПЭ перейти к следующему этапу дискуссий, который мог бы быть сфокусирован на проработке таких элементов, как определения и другие концепции с потенциалом сужения спектра рассматриваемых предметов до стратегических направлений.
- 3. В целях стимулирования дискуссий и оказания помощи делегациям в подготовке к работе ГПЭ Председатель предлагает следующие вопросы, сведенные в три широкие категории. Эти вопросы призваны лишь помочь начать субстантивный обзор каждой из трех категорий. Они не являются исчерпываю-

GE.17-15299 (R) 280917 280917





щими и не представляют в полной мере все мнения по данному предмету; в частности, некоторые вопросы директивного плана, которые относятся к следующему этапу дискуссий, были пока оставлены в стороне.

#### Технологии

- Какие технологии вносят или могли бы вносить свой вклад в смертоносную автономность оружейных систем? Можно ли их в общем плане охарактеризовать как искусственный интеллект/автономные системы (ИИ/АС)?
- Для каких гражданских целей они используются или могут использоваться?
- Являются ли наиболее наглядными примерами автономных систем физические роботы (обособленные элементы) или виртуальные машины (расширенные системы обработки информации)? Возможны ли альтернативные технологические пути (например, нейроморфные, усиление человеческих возможностей), которое не охвачены дискуссией по АСОЛД?
- Какие различные аспекты автономности усматриваются в нынешнем наборе технологий ИИ/АС? На каком уровне мы находимся в этих аспектах (энергетическая автономность, вычислительная автономность и т.д.) по сравнению с полной автономностью (даже в идеальном варианте)?
- Имел ли место в последние годы (созданная в «Google» сеть, способная распознавать изображения кошек, и разработанный компанией «Deep Mind» алгоритм DQN для двухмерных игр игровой консоли «Atari 2600», виртуальный собеседник «Женя Густман», прошедший тест Тью-ринга, программа «AlphaGo» и т.д.) или может ли произойти переход от конкретных приложений (узкий или слабый ИИ) к общему искусствен-ному интеллекту (ОИИ), следствием чего может стать «интеллектуальный взрыв»/«технологическая сингулярность»/появление искусственного «суперинтеллекта» (ИСИ)?
- Каким образом существующие системы подвергаются проверке (правильно ли она была построена) и аттестации (была ли построена правильная система)? Поддаются ли существующие и проектируемые автономные системы обследованию (что ты знаешь и откуда ты это знаешь)? Могут ли машины описать свое обучение?
- Каким образом промышленность решает проблемы безопасности человека и общественной безопасности (например, хакерство и невмешательство в частную жизнь)? Могут ли автономные машины быть защищены от хакерского воздействия?
- Возможно ли существование программных/аппаратных средств, блокирующих определенные схемы поведения машины, и можно ли воспрепятствовать тому, чтобы обучающаяся машина могла обойти/изменить их?
- Ограничивает ли каким-либо образом дискуссию по АСОЛД трансформирующийся характер ИИ и его возможная вездесущность, или же ИИ сходен с другими технологиями двойного применения, которые появлялись в прошлом?

#### Военное воздействие

• В каких конкретных областях применяются или могут применяться автономные технологии? Материально-техническое обеспечение? Разведка, наблюдение и рекогносцировка? Охрана границы и оборона района? Может характер их применения привести к постепенному изменению задач, которые ставятся для АСОЛД, или иным последствиям?

**2** GE.17-15299

- Имеются ли определенные сферы (например, надводная и подводная), где потенциальное применение АСОЛД является более вероятным, и почему?
- На какой наступательный или оборонительный потенциал АСОЛД оказывали бы наибольшее воздействие? Может ли потенциальное применение АСОЛД снизить порог применения силы? Может ли оно расширить масштабы асимметричного наращивания сил или скрытого применения силы? Являются ли соображения, связанные с гонкой вооружений и поддержанием стабильности, актуальными для КНО?
- Могут ли потенциальные АСОЛД распространяться/обучаться для действий совместно с террористами и другими незаконными негосударственными субъектами?
- Имеются ли определенные виды военного применения (удаление взрывоопасных пережитков войны, разминирование и т.д.), которые могли бы быть желательными даже в нынешнем контексте КНО?
- Могут ли потенциальные АСОЛД быть интегрированы в существующие системы военного командования и контроля? Имеются ли определенные доктринальные проблемы, связанные исключительно с АСОЛД?
- Как потенциальные АСОЛД согласуются с существующими концепциями транспарентности в вооружениях в плане развертывания, обладания, торговли или передач?

## Правовые/этические вопросы

- На ком лежит юридическая и судебная ответственность за существующие или проектируемые автономные системы? На проектировщике-разработчике, законном собственнике, пользователе и/или машине?
- Каковы основные черты национальных или региональных законов, планируемых или уже введенных в действие для регулирования автономных систем, таких как автомобили без водителя, роботы, предназначенные для развлечения, или роботы-компаньоны и виртуальные собеседники?<sup>1</sup>
- Существуют ли кодексы поведения или отраслевые стандарты, которые могут быть актуальными для дискуссии по АСОЛД?<sup>2</sup> Что актуально для КНО в этих опорных документах?
- Имеются ли регламентационные ориентиры из других областей, таких как генетика?
- Могут ли нормы международного гуманитарного права, разработанные для человеческого поведения и поведения, контролируемого государством, продолжать применяться mutatis mutandis к потенциально автономным машинам и через какие посреднические механизмы?
- Как потенциальные АСОЛД согласуются с существующими национальными обзорами законности оружейных систем, проводимыми до начала разработки, тестирования и развертывания?

GE.17-15299 3

<sup>&</sup>lt;sup>1</sup> Например, закон Республики Корея «О разработке и распространении интеллектуальных роботов» и доклад Комитета Европейского парламента по правовым вопросам (2015/2103(INL)).

<sup>&</sup>lt;sup>2</sup> Такие как Р7000-7002 Института инженеров электротехники и электроники (ИЕЕЕ) в рамках Глобальной инициативы по этическим аспектам в сфере искусственного интеллекта и автономных систем и Принципы робототехники Совета по инженерным и физическим научным исследованиям и Совета по гуманитарным научным исследованиям Соединенного Королевства.

- Имеются ли правовые лакуны в отношении потенциальных автономных систем в других областях международного регулирования, таких как торговля (правила происхождения и т.д.) и интеллектуальная собственность?
- Могут ли машины быть действительно интеллектуальными по аналогии с людьми (феноменально сознательные, руководствующиеся соответствующими намерениями, творческие, способные к сопереживанию, эволюционирующие, свободные агенты с интегрированным интеллектом)?
- Какая система этических норм или конкретных этических ценностей может применяться к машинам? Каким образом различные взгляды на сознание и взаимодействие между человеком и материальным миром, существующие в различных этических/теологических системах, могут быть взаимно согласованы в универсально применимом наборе запретов для обучающихся машин?
- Озабоченности, связанные с этическими/нравственными аспектами, до сих пор фокусировались на том, что машины будут лишать людей жизни. Как следует относиться к совместным действиям человека и машины или к усилению человеческих возможностей?
- Как могут кодироваться эти этические запреты? Может ли машина стать нравственным существом?

**4** GE.17-15299