

Противоречат ли автономные (интеллектуальные) технические системы военного назначения международному праву?

Огородов Дмитрий Владимирович
член Российской ассоциации
искусственного интеллекта,
член рабочей группы по этике IEEE,
к.ю.н., арбитр МКАС при ТПП РФ

d.ogorodov@gmail.com

Дилетантизм и «хайп» – главная проблема правовой регламентации технологий искусственного интеллекта

- ❑ Сленг маркетологов и управленцев («цифровая экономика» = «цифровые права») стал подменять категориальный аппарат юриспруденции (система права и методы регулирования; правоотношения – их субъекты, объекты и содержание, юридические факты и др.). Сомнительная новелла ст. 141.1 «Цифровые права» ГК РФ;
- ❑ Незнание отечественных (СССР, РФ) достижений в области технологий ИИ. Подмена научной дискуссии маркетингом готовых технических решений иностранного происхождения. Создание вчерашним студентом несложной нейронной сети **в готовом программном пакете** Neural Network Toolbox среды разработки MATLAB или с помощью библиотеки Keras на языке Python **подается как достижение**;
- ❑ Некритичное восприятие юристами **фейковых новостей** на тему ИИ: «робот София получил гражданство Саудовской Аравии» – первая ссылка в списке литературы, которой д.ю.н. М.Рожкова обосновывает научную актуальность темы https://zakon.ru/blog/2019/11/23/iskusstvennyj_intellekt_i_intellektualnye_roboty__что_это_такое_или_кто_это_такie;
- ❑ Основанные на некритичном цитировании иностранных публикаций диссертации по юридическим наукам (докторская диссертация П.М.Морхата, кандидатская диссертация Е.П.Сесицкого). Научообразный текст этих диссертаций заставляет предполагать, что техническую сторону дела диссертанты изучали посредством 15-минутного «гугления»;
- ❑ Создание «сильного ИИ» провозглашено в числе целей Национальной стратегии развития искусственного интеллекта на период до 2030 года (утв. Указом Президента РФ от 10.10.2019 № 490). **«Сильный ИИ» – гипотеза американского лингвиста и философа Дж.Сёрла**, никогда не занимавшегося созданием реальных автономных / интеллектуальных систем;
- ❑ российская государственная Стратегия развития искусственного интеллекта готовилась кулуарно отдельными игроками крупного бизнеса, без широкой научной дискуссии, отчего Стратегия больше напоминает бизнес-план ПАО «Сбербанк», нежели общегосударственную программу развития

«Смертоносные автономные системы» и «роботы-убийцы» – политически ангажированная неюридическая риторика журналистов и политологов

- Введение в заблуждение публики со стороны гиперактивных журналистов (Алена Попова) и политологов (В. Козюлин), **не обладающими познаниями** ни в области права, ни в области разработки и применения вооружения и военной техники.
Пример: выступление Поповой перед философами на семинаре Научного совета по методологии искусственного интеллекта РАН в мае 2019
- Продвижение политических инициатив непонятных иностранных организаций;
- Психологическое манипулирование:
 - ссылка на авторитеты (*«мы обсуждаем на площадке ООН», «за нами – нобелевские лауреаты и медийные бизнесмены»*);
 - использование заведомо негативных в русском языке выражений *«роботы-убийцы», «смертоносные», «убийство»*.
Однако *to kill* применительно к оружию не тождественно глаголу «убивать» в смысле ст. 105 УК РФ. Применительно к оружию (средствам поражения) и ведению боевых действий вооруженными силами *to kill* означает «поражать живую силу», «наносить летальные поражения». Ведущая боевые действия мотострелковая рота не является организованной группой лиц, совершающих убийство (п.«ж» ч. 2 ст. 105 УК РФ). Оборонительная война для отечественной культуры традиционно правомерна и этически одобряема;
- Неудачная терминология. Фонетическая калька с английского языка *«lethal autonomous weapon»*. Однако *«нелетальное оружие»* **не является парной категорией**, выражение *«летальное оружие»* - плеоназм.

«Смертоносные автономные системы» и «роботы-убийцы» – политически ангажированная неюридическая риторика журналистов и политологов

Отсутствие **юридической аргументации** у журналистов и политологов (типовой диалог):

- *Роботы-убийцы смертоносны, они противоречат международному праву. Мы должны запретить роботов-убийц!*
- *И каким же нормам противоречат, позвольте поинтересоваться*
- *Статье 36 первого протокола и оговорке Мартенса!*
- *?!*

При изучении, разработке, приобретении или принятии на вооружение **новых видов оружия**, средств или **методов ведения войны** Высокая Договаривающаяся Сторона **должна определить, подпадает ли** их применение, при некоторых или при всех обстоятельствах, **под запрещения**, содержащиеся в настоящем Протоколе или в каких-либо других нормах международного права, применяемых к Высокой Договаривающейся Стороне.

статья 36 «Новые виды оружия» Дополнительного протокола (Протокола I) к Женевским конвенциям от 12 августа 1949 года, касающийся защиты жертв международных вооруженных конфликтов

«Три группы угроз смертоносных автономных систем»: аргументы В. Козюлина

За последнее десятилетие в мире ускорилось появление технологий, способных придать традиционным вооружениям уникальные возможности, присущие боевым роботам. На площадке ООН для таких вооружений определили термин — смертоносные автономные системы (САС). Так называют **оружие, способное поражать цели на земле, в воздухе, на воде и под водой без участия человека.**

Автономные ударные системы с использованием искусственного интеллекта создают угрозы, которые можно разделить на три группы:

Первая группа включает риски, связанные с исключением человека из цикла принятия решений на применение оружия — так называемая «проблема значимого человеческого контроля».

Представители мировой общественности (неправительственные организации «Stop Killer Robots», «Article 36», «International Committee for Robot Arms Control», бизнесмены и ученые, в частности, **Стивен Хокинг, Илон Маск, Стив Возняк**) считают, что полностью автономное оружие с большой вероятностью **не будет в состоянии соответствовать международному гуманитарному праву и правам человека**, создаст проблему с определением ответственных лиц в случае противоправных действий автономных аппаратов.

Боевым **роботам ставится в вину неспособность испытывать сочувствие**, то есть то человеческое чувство, которое нередко выступает сдерживающим фактором в отношении применения оружия.

Еще один аргумент против САС — их использование противоречит принципу гуманности и требованиям общественного сознания.

<https://russiancouncil.ru/analytics-and-comments/analytics/tri-gruppy-ugroz-smertonosnykh-avtonomnykh-sistem/>

Вадим Козюлин, к.полит.н., профессор Академии военных наук, директор проекта по новым технологиям и международной безопасности ПИР-Центра, эксперт РСМД

«Три группы угроз смертоносных автономных систем»: аргументы В.Козюлина

Вторая группа угроз связана с риском нарушения стратегической стабильности. **Элементы автономности и искусственного интеллекта приходят во все сферы военного противостояния.** Так, в ядерной сфере появляются высокоточные тактические ядерные бомбы и гиперзвуковые аппараты с новыми ядерными боеголовками; в **космической области — беспилотные космические дроны**, системы низкоорбитальных спутников наблюдения и связи; в сфере ПРО — связка системы слежения и наблюдения с системой управления и связи; в киберсфере — кибероружие и **автоматизированная система ответного киберудара**. Некоторые из этих видов вооружений, как, например, гиперзвуковые ракеты или средства кибернападения, могут служить инструментом стратегического сдерживания наряду с ядерным оружием. То есть даже у неядерных стран возникает возможность резкого наращивания как потенциала сдерживания, так и ударных возможностей. За этими тенденциями стоит серия рисков:

- риск установления технологического и военного глобального превосходства одной из стран;
- новая гонка вооружений;
- повышение региональной и международной напряженности;
- сокращение транспарентности военных программ;
- **игнорирование международного права;**
- распространение опасных технологий среди негосударственных субъектов.

Основываясь на опыте военного использования военных и коммерческих дистанционно управляемых беспилотных летательных аппаратов (БПЛА), исследователи делают вывод о том, что технологии производства САС, **их компоненты и программное обеспечение получают значительное распространение в мире**, что станет толчком для гонки вооружений, **приводящей к усилению нестабильности** и эскалации разнообразных рисков.

Некоторые эксперты полагают, что поддержание стратегической стабильности в ближайшие десятилетия потребует пересмотра основ теории сдерживания в многополярном мире»

«Три группы угроз смертоносных автономных систем»: аргументы В.Козюлина

Третья группа угроз вытекает из радикального сокращения времени на принятие стратегических решений в рамках системы сбора и анализа разведывательной информации (Intelligence, Surveillance & Reconnaissance (ISR)) и военного управления и связи (Communications, Command and Control (C3)). Главный недостаток человека в сравнении с машиной состоит в том, что человеческому мозгу необходимо слишком много времени для оценки ситуации и принятия верного решения. Целый ряд военных программ в ведущих государствах мира (в частности программы Пентагона Maven, COMPASS, Diamond Shield) нацелен на то, чтобы суперкомпьютеры взяли на себя работу по анализу различных данных и подготовке сценариев для политического и военного руководства.

С этим связаны, как минимум, такие риски:

- **нехватка времени для принятия осмысленных решений.**
- недостаточный контроль над ситуацией со стороны человека.
- принятие стратегических решений на основе математических алгоритмов и систем машинного обучения, а не человеческой логики.
- отсутствие взаимопонимания между машиной и человеком. **Нейронные сети пока не в состоянии объяснить закономерности своей работы на человеческом языке.**

<https://russiancouncil.ru/analytics-and-comments/analytics/tri-gruppy-ugroz-smertonosnykh-avtonomnykh-sistem/>

Вадим Козюлин, к.полит.н., профессор Академии военных наук, директор проекта по новым технологиям и международной безопасности ПИР-Центра, эксперт РСМД

Автономные (интеллектуальные) системы военного назначения: этика и международное право

Дискуссионные вопросы:

- ❖ являются ли военные АИС и применение их вооруженными силами **новыми средствами и методами ведения войны**?
- ❖ замена людских сражений **безлюдными сражениями роботов аморальна?** Или наоборот, гуманизирует войну?
- ❖ боевые АИС сокращают или увеличивают жертвы среди некомбатантов и их страдания? Вопрос требует детального обсуждения. Между тем сторонники запрета **не имеют статистики** и другой развернутой военно-технической аргументации по обсуждаемому вопросу. Например, нет открытых исследований по вопросу ошибок распознавания целей системами самонаведения с элементами ИИ.
Аналогии из гражданской сферы (системы медицинской диагностики с элементами ИИ) показывают неоднозначность сравнения эффективности человеческого и машинного распознавания;
- ❖ **мнимая проблема ответственности:** АИС не являются и не могут являться субъектами права, а потому **не могут освобождать от ответственности** военнослужащих и их военно-политических руководителей за военные преступления;
- ❖ сторонники запрета **не могут соотнести свою терминологию** («оружие, способное поражать цели на земле, в воздухе, на воде и под водой без участия человека») с **реально существующими системами оружия**. Например, под это определение попадает минное оружие, известное человечеству без малого 200 лет;
- ❖ сторонники запрета не могут предложить эффективные способы контроля соблюдения международных запретов, учитывая деликатность сферы регулирования. Никакое государство никого не подпустит к программному обеспечению оружия.

Автономные (интеллектуальные) системы военного назначения: этика и международное право

Вопросы этики и международной правовой политики в контексте интересов России:

- ❖ **этическое обоснование запрета «спуска курка» роботом**: бесспорность существования или сомнительность политических игр?
- ❖ морален **или аморален запрет военных АИС для РФ** со стороны, например, Австралии? Учитывая разницу людских потерь СССР и Австралии в ВМВ, географического положения и протяженности границ РФ, численности населения РФ и др. государств)
- ❖ дестабилизирует **или стабилизирует миропорядок**, демотивируя потенциального агрессора, полумифическая автономная система ответного ядерного удара «Периметр»?

ВЫВОДЫ:

- ❖ на сегодня не представлено убедительных правовых и этических (политико-правовых) доводов в пользу запрета автономных (интеллектуальных) систем военного назначения;
- ❖ de lege lata разработка и применение АИС военного назначения подчиняется общим нормам международного права вооруженных конфликтов;
- ❖ с научно-правовой точки зрения можно выделить **гипотетическую проблему** международно-правового запрета или ограничения разработки технологий искусственного интеллекта, **способных выйти из-под контроля человечества** и создать угрозу его существованию (агрессивные самореплицирующиеся роботы, *SkyNET*) – при условии, что такой технологии будут когда-либо созданы.

Спасибо за внимание!