Научная статья
УДК 343.01:004
DOI: 10.47475/2311-696X-2024-42-3-12-19

С. 12–19

«ВИНА» ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ЗА «ОШИБКУ» И ПРОБЛЕМА ЯДЕРНОЙ ВОЕННОЙ БЕЗОПАСНОСТИ В КОНТЕКСТЕ ЮРИДИЧЕСКОЙ ТОПИКИ

Юрий Михайлович Батурин
Московский государственный университет имени М. В. Ломоносова, Москва, Россия
baturin@ihst.ru
https://orcid.org/0000-0003-1481-5369

Аннотация. Обсуждение в рабочей группе ООН проекта конвенции по использованию информационно-коммуникационных технологий в военно-политических целях еще недавно давало определённую надежду на компромисс, но сейчас очевидно, что позиции основных сторон — США и России — непримиримы. Особенно деликатны вопросы о кибербезопасности военных систем. Юридическая наука может предложить нестандартные методы для разрешения этой коллизионной ситуации. Среди возможных инструментов — теория мер безопасности как правового института, включающего санкции безопасности, разработанная известным специалистом по уголовному праву профессором Н. В. Щедриным.

В последнее время актуальным (и спорным) стал вопрос об использовании для различных прикладных целей «искусственного интеллекта». Вокруг него возникло множество мифов. Для юридически корректного решения возникающих проблем в уголовном, гражданском и других отраслях права требуется демистификация «искусственного интеллекта». В условиях недостатка информации о реальном применении «искусственного интеллекта» повышенное беспокойство стали вызывать системы управления ядерным оружием вероятного противника, к которым причисляют комплексы автоматического управления ответным ядерным ударом СССР/России и США. Предлагается нестандартный юридический подход к выработке проекта конвенции, снижающий уровни риска и недоверия к «интеллектуальным» системам управления ядерным оружием.

Ключевые слова: международная информационная безопасность, кибербезопасность, информационно-коммуникационные технологии, «искусственный интеллект», «ошибка» и «вина» искусственного интеллекта, смертоносные автономные системы вооружений, системы управления ядерным оружием, меры безопасности как правовой институт, юридическая топика, экспериментальная юриспруденция, «ядерная зима», моделирование

Для цитирования: Батурин Ю. М. «Вина» искусственного интеллекта за «ошибку» и проблема ядерной военной безопасности в контексте юридической топики // Правопорядок: история, теория, практика. 2024. № 3 (42). С. 12–19. DOI: 10.47475/2311-696X-2024-42-3-12-19

Research article

ARTIFICIAL INTELLIGENCE “FAULT” FOR “ERROR” AND THE PROBLEM OF NUCLEAR MILITARY SECURITY IN THE CONTEXT OF A LEGAL TOPICALITY

Yuri M. Baturin
Lomonosov Moscow State University, Moscow, Russia
baturin@ihst.ru
https://orcid.org/0000-0003-1481-5369

Abstract. The discussion in the UN Working Group on the draft convention on the use of information and communication technologies for military and political purposes has recently given some hope for a compromise, but now it is obvious that the positions of the main participants — the United States and Russia — are irreconcilable. Questions about the cybersecurity of military systems are particularly delicate. Legal science can offer non-standard methods to resolve this conflict situation. Among the possible tools is the theory of security measures as a legal institution that includes security sanctions, developed by the well-known criminal law specialist Professor N. V. Shchedrin.

Recently, the issue of using "artificial intelligence" for various applied purposes has become topical (and controversial). Many myths have arisen around it. For legally correct solution of arising problems in criminal, civil and other branches of law, demystification of "artificial intelligence" is required. Given the lack of information about the real use of "artificial intelligence", the systems of control of nuclear weapons of a probable enemy, which include the complexes of automatic control of retaliatory nuclear weapons of the USSR/Russia and the United States, have become a source of increased concern. The article proposes a non-standard legal approach to drafting a convention that would reduce the levels of risk and mistrust of "intelligent" nuclear weapons control systems.

Keywords: international information security, cybersecurity, information and communication technologies, "artificial intelligence", "error" and "guilt" of artificial intelligence, lethal autonomous weapon systems, nuclear weapon control systems, security measures as a legal institution, legal topics, experimental jurisprudence, "nuclear winter", modeling

For citation: Baturin YuM. "Guilt" of artificial intelligence for "error" and the problem of nuclear military security in the context of legal topics. *Pravoporyadok: istoriya, teoriya, praktika* [Legal and Order: History, Theory, Practice]. 2024;(3):12-19. (In Russ.) DOI: 10.47475/2311-696X-2024-42-3-12-19

Введение

Обсуждение документов по международной информационной безопасности в ООН еще несколько лет назад давало определённую надежду на компромисс, но сейчас очевидно, что позиции основных сторон переговоров (США и России) непримиримы. Особенно деликатны вопросы информационной безопасности военных систем. В последнее время актуальным (и спорным) стал вопрос об использовании для различных прикладных целей «искусственного интеллекта». Вокруг него возникло множество мифов. Для юридически корректного решения возникающих проблем в уголовном, международном и других отраслях права требуется демистификация «искусственного интеллекта». В условиях недостатка информации о реальном военном применении т. н. «искусственного интеллекта» беспокойство стали вызывать системы управления ядерным оружием вероятного противника, к которым причисляют комплексы автоматического управления ответным ядерным ударом СССР/России и США.

Методы исследования

Использованы нормы международного военного и гуманитарного права, информационного права, включая стратегические документы по искусственному интеллекту, а также нестандартные методы, предлагаемые юридической наукой для разрешения коллизионной ситуации, сложившейся в сфере международной информационной безопасности, такие как теория мер безопасности как правового института, включающего санкции безопасности, разработанная известным специалистом по уголовному праву профессором Н. В. Щедриным [1, с. 529–608], юридическая топика [2] и экспериментальная юриспруденция [3, с. 27–35].

Терминология

На официальном уровне Россия использует в документах и в ходе дипломатических переговоров понятие «международная информационная безопасность», подразумевающее как технические, так и гуманитарные угрозы. США с союзниками предпочитают понятие «кибербезопасность», которое несколько уже, поскольку редко выходит за пределы угроз технических. В документах ООН в качестве компромиссного варианта применяется понятие «безопасность в сфере информационно-коммуникационных технологий (далее — ИКТ) и их использования». В отечественных научных работах встречаются все три понятия. Представляется предпочтительным использование, как более точно отражающего суть дела на русском языке, понятия «международная информационная безопасность». Термин «кибербезопасность» (и производные «киберпространство», «киберпреступление» и др.) пришел в мир из англоязычной художественной литературы (Gibson W. Burning Chrome 1982; Neuromancer 1984) и превращался в понятие не столько научным путем, сколько спонтанно, в зависимости от того, чем его наполняли авторы (если и когда чем-то наполняли). Понятие, принятое в ООН, слишком длинное для частого употребления. В статье все три понятия будут использоваться как синонимы (хотя это не совсем корректно), но строго в соответствии с контекстом.

В Соединенных Штатах в последние годы активно используется понятие «новые разрушительные технологии», в которые включают «автоматические системы принятия решений», а сами автоматические системы разделяют на несколько категорий по степени участия человека. В советских/российских научно-технических традициях системы управления классифицируются по критерию участия/неучастия человека. В автоматизированных системах решения принимаются человеком, включенным в контур управления («man in the loop»). Системы автоматического управления действуют строго по алгоритмам, заложенным человеком, но не включенным в контур управления («man out of the loop»). Все перечисленные системы, кроме автоматических, могут быть объединены введенным в юридический оборот понятием «автономные системы управления оружием», которые Международный комитет Красного Креста определяет как «любые системы вооружений с автономией в ее важнейших функциях. То есть система управления оружием способна отбирать (то есть искать или обнаруживать, идентифицировать, отслеживать, выбирать) и атаковать (то есть использовать силу, нейтрализовать, разрушать или уничтожать) цели без непосредственного

человеческого вмешательства»[1]. Казалось бы, должно быть наоборот: к автономным следует отнести именно автоматические системы управления. Но это не так. Тонкость в определении: «…с автономией в ее важнейших функциях». Не абсолютная автономия, а только автономия некоторых функций. Ситуацию проясняет параграф 2 того же документа: «Автономность в существующем оружии. После активации человеком-оператором система оружия с помощью своих сенсоров и компьютерных программ выбирает цель и начинает атаку»[2]. То есть автономность начинается *после* активации человеком-оператором.

Необходимо сделать оговорку и о понятии «искусственный интеллект». Словосочетание «Artificial Intelligence», введенное в 1956 г. Дж. МакКарти на семинаре по информатике в Дартмутском колледже (США) означало искусственную систему, способную разумно рассуждать, но было неправильно переведено на русский язык как «Искусственный интеллект» (ИИ). Для слова «интеллект» в английском языке есть другое слово — «Intellect». В результате возникло мифологизированное понятие со всеми недоразумениями и ошибочными выводами. В данной статье «искусственный интеллект» будет означать информационную систему, способную, сравнивая показания большого количества датчиков, логически выводить и рекомендовать человеку решение. В такой формулировке ИИ может в разной степени использоваться во всех автономных системах управления — и в автоматизированных, и в автоматических, при оговорке, что пока такие системы с действительно полноценным искусственным интеллектом пока не созданы, хоть и их создание, по оценкам экспертов, весьма вероятно в перспективе 30–40 лет.

Предмет анализа

В документе «Принципы, касающиеся международной информационной безопасности», предложенном Россией 12 мая 1999 года на рассмотрение Генеральной ассамблеи ООН (далее — ГА ООН), содержалось определение: «Международная информационная безопасность — состояние международных отношений, исключающее нарушение мировой стабильности и создание угрозы безопасности государств и мирового сообщества в информационном пространстве». 23 сентября 1998 года министр иностранных дел РФ И. С. Иванов направил генеральному секретарю ООН К. Аннану специальное послание, в котором содержался проект резолюции ООН «Достижения в сфере информатизации и телекоммуникаций в контексте международной безопасности» (A/RES/53/70)[3]. В нем впервые была закреплена «триада угроз» в сфере международной информационной безопасности:

— использование ИКТ в военно-политических целях (ущемление суверенитета, создание военной угрозы и пр.);

— использование ИКТ в преступных целях (неправомерное использование информации, распространение вредоносных программ и пр.);

— использование ИКТ в террористических целях (включая деятельность в социальных сетях).

В 1999 г. резолюция была обновлена (A/RES/54/49)[4], в ней впервые зафиксирована возможность отрицательного воздействия ИКТ на безопасность в военной сфере [4, с. 6–12]. В связи со стремительной милитаризацией информационного пространства наибольшую опасность сегодня представляет группа военно-политических угроз, а в ней — использование информационных технологий в системах вооружений. Минимальный уровень доверия России и США друг к другу связан с автономными военными системами управления ядерным оружием, оснащенными так называемым «искусственным интеллектом». Именно этой проблеме посвящена статья.

История переговоров

56-я ГА ООН в 2001 году по инициативе России создала Группу правительственных экспертов (далее — ГПЭ, Группа) из 15 государств-членов ООН. Работа ГПЭ началась в 2004 году. Согласовать доклад ГПЭ не удалось из-за того, что США выступили против включения в документ тезисов о военно-политическом измерении угроз МКТ, не желая брать на себя международные обязательства в военной сфере. По двум другим направлениям США сотрудничество продолжали. ГПЭ возобновила свою работу в 2009 году, и в 2010 году Группе удалось согласовать доклад, содержащий, в основном, вопросы терминологии. В 2013 году ГПЭ третьего созыва согласовала доклад, не исключающий выработку в будущем дополнительных норм международного права для регулирования деятельности государств в ИКТ-среде. В 2014 году состав ГПЭ увеличили до 20 государств. Обсуждение проблемы выявило острые противоречия, тем не менее в 2015 году был принят итоговый доклад, в котором, в частности, подтверждалась возможность разработки новых правовых норм в области информационных технологий. Только в 2018 году переговорный процесс возобновился. В 2019 создается рабочая группа открытого состава (далее — РГОС) из представителей 193 государств, но общую позицию выработать не удалось — появились два проекта — РФ и США.

В 2021 году в РГОС впервые в истории переговорного процесса удалось вывести делегацию США

[1] Views of the International Committee of the Red Cross (ICRC) on autonomous weapon system : 11 April 2016. [Par. 1] // ICRC. Convention on Certain Conventional Weapons (CCW) : Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) : 11–15 April 2016, Geneva. URL: https://www.icrc.org/en/download/file/21606/ccw-autonomous-weapons-icrc-april-2016.pdf (дата обращения: 15.04.2024).

[2] Views of the International Committee of the Red Cross (ICRC) on autonomous weapon system : 11 April 2016. [Par. 1].

[3] URL: https://www.securitycouncilreport.org/atf/cf/%7B65BFCF9B-6D27-4E9C-8CD3-CF6E4FF96FF9%7D/a_res_53_70.pdf (дата обращения: 15.04.2024).

[4] URL: https://ifap.ru/ofdocs/un/5449.pdf (дата обращения: 15.04.2024).

на признание в тексте резолюции ГА ООН возможности выработки дополнительных договоренностей по безопасности в ИКТ-среде, имеющих обязательную силу. В 2022 году в резолюции ГА ООН 73/27[1] удалось сформулировать 13 правил, норм и принципов ответственного поведения государств в ИКТ-среде и признать их источником мягкого права. Но в дальнейшем США свернула на путь легализации односторонних действий, например, через концепцию коллективной атрибуции, декларирующей возможность для группы государств без предъявления доказательств осуществлять превентивные удары в ответ на якобы имевшую место информационную агрессию [4, с. 13–22, 38–42].

В настоящее время позиции РФ и США несовместимы и непримиримы, перспективы договориться призрачны. Более того, директор департамента международной информационной безопасности МИД РФ Артур Люкманов заявил 10 апреля 2024 года: «Цель США в сфере ИКТ — нанести нам поражение с помощью некинетических средств, то есть с помощью ИКТ. Тут невозможно ни о чём договариваться»[2].

Применение юридической топики

Топика со времен Аристотеля опирается на понятие «топ», «топос» (*др.-греч. τόπος*) — место, общее место не в смысле банальности, а как смысловая общепонятность. Юридическая топика — способ проблемно-ориентированного мышления, не ограниченного рамками конкретных правовых концепций, основанного на поиске общих мест и заключающийся в поиске взаимоубедительных аргументов и взаимоприемлемых решений. Топическое мышление позволяет посредством интерпретации открыть новые варианты понимания без разрушения старых [2].

Переход от зашедшего в тупик классического переговорного процесса в отношении «интеллектуальных» военных систем к топическим переговорам предполагает перенесение акцента с согласования принципов и формулирования соответствующих договорных норм на выявление наиболее чувствительного общего места, юридически аутентичное и семантически адекватное его описание на русском и английском языках и последующее последовательное расширение этого общего места. Можно рекомендовать РФ и США строить переговорную аргументацию, удерживая ее логическую последовательность в рамках здравого смысла. При этом вовсе не обязательно сразу же отказываться от своих «высших принципов». Достаточно на первое время оставить их за «красными линиями», которые не переступать в ходе переговоров. А начать требуется с выявления топа (общего) — с чем согласны обе стороны.

Сейчас общим местом для РФ И США являются: а) отсутствие какой-либо договоренности по автономным системам управления ядерным оружием и б) установленные в 2023 году «Часы Судного дня» (*англ. «Doomsday Clock»*) за 1 мин 30 сек до полуночи, символизирующей ядерную катастрофу. Журнал Чикагского университета «Бюллетень учёных-атомщиков» в 1947 году предложил такой наглядный образ: колебания минутной стрелки показывают, насколько мир отстоит от ядерной войны. В 1947 году стрелку поставили на 7 минут до полуночи. За 76 лет минутную стрелку отводили назад 8 раз, в том числе однажды на максимальное расстояние (17 минут до полуночи) — в 1991 году, когда между СССР и США был подписан договор о сокращении стратегического вооружения. Ближе, чем сегодня (полторы минуты до полуночи) стрелки часов никогда не находились. Переговоры по этому единственному вопросу могут принести результат и тем быстрее, чем дальше они будут вестись от «красных линий» (принципов безопасности в ИКТ-среде). Достигнутое в случае успеха соглашение, при всей скромности охвата им всех проблем международной информационной безопасности, между тем, могло бы стать исходной площадкой для более масштабных переговоров по международной безопасности в ИКТ-среде.

Чуть ли не единственным *местом* в прямом смысле слова (консультативной площадкой), где обсуждаются сложные проблемы, подобные описываемой, остался дистанционно проводимый российско-американский диалог (Committee on International Security and Arms Control — CISAC), созданный в 1980 году Национальной академией наук (NAS) США. Российской стороной диалога выступает Институт США и Канады имени академика Г. А. Арбатова РАН. От России председательствует член Российского Пагуошского комитета при Президиуме РАН академик РАН С. М. Рогов[3]. С обеих сторон участвуют эксперты — дипломаты и военные высокого ранга в отставке и ученые.

«Искусственный интеллект» для «Мертвой руки»

Наибольшее недоверие у американских военных вызывает созданный ещё в Советском Союзе в 1985 году (что отнюдь не исключает его модернизацию) комплекс автоматического управления массированным ответным ядерным ударом «Барьер» (позднее — «Периметр»), с легкой руки журналистов известный на Западе как «Мёртвая рука» («Dead Hand»), задача которого — гарантированно донести до командных пунктов и отдельных пусковых установок приказ о пуске. В США и НАТО рассматривают ее как полностью автономную и недоверчиво заявляют, что реальное назначение системы неизвестно. Заметим, что задолго до появления «Периметра» в США в 1961 году создали дублирующую систему управления стратегическим ядерным оружием, названную «Операция Зазеркалье» («Operation Looking Glass»).

[1] URL: https://documents.un.org/doc/undoc/gen/n18/418/07/pdf/n1841807.pdf (дата обращения: 15.04.2024).

[2] DigitalRussia (Цифровая Россия). URL: https://t.me/drussia/26482 (дата обращения 15.04.2024).

[3] См.: Российские и американские учёные обсудили приостановку участия России в Договоре СНВ-3 // Российская академия наук. URL: https://www.ras.ru/news/shownews.aspx?id=d28ae97c-b2d8-44f0-88db-9f6d0cb23c67 (дата обращения: 15.04.2024).

Позже в США был создан прямой аналог советского «Периметра» — комплекс передачи приказа о запуске ракет (AN/DRC-8), названный «Система экстренной связи для ракетного пуска» («Emergency Rocket Communications System»). Таким образом, и у российской стороны тоже имеются основания для недоверия. В атмосфере взаимной подозрительности, предубеждения и недостаточного технического понимания работы автоматических систем управления, на фоне с избытком мифологизированного «искусственного интеллекта» с наносными слоями ставшей модной темой «вины» ИИ за «ошибку», проблема автономной системы управления ядерным оружием стала более опасна, чем была сама по себе в «чистом виде».

Что такое ИИ и его «вина» за свою «ошибку»?

Определение «искусственного интеллекта» возьмем из «Национальной стратегии развития искусственного интеллекта на период до 2030 года», утвержденной Указом Президента РФ № 490 «О развитии искусственного интеллекта в Российской Федерации» от 10.10.2019[1]. Это определение было подкорректировано Указом Президента Российской Федерации от 15 февраля 2024 г. № 124: «Искусственный интеллект — комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений» (подп. „а“ п. 5)[2]. Обратим внимание на то, что из предыдущей версии определения изъято «самообучение» ИИ, и, с другой стороны, отдельно определен «сильный искусственный интеллект — тип искусственного интеллекта, который способен выполнять различные задачи, взаимодействовать с человеком и самостоятельно (без участия человека) адаптироваться к изменяющимся условиям (подп. „х“ п. 5)[3]. «Сильный ИИ» сможет самообучаться, и задачу его создания скорректированная стратегия ставит на дальнюю перспективу. Это технический аргумент за то, что пока автономной системы управления ядерным оружием у РФ нет. Но политикам нужны иные аргументы.

Юристов с первого дня обучения в бакалавриате учат «думать как юрист»: по сути, рассуждать в одном деле по аналогии к другим. В частности, в контексте стремительно меняющихся технологий закон почти всегда рассматривает новую технологию как новую форму прежней. От аналогии один шаг до метафоры. По определению, идущему еще от Аристотеля, метафора (*др.-греч. μεταφορά — «перенос», «переносное значение»*) — свернутое (неназванное) сравнение одного объекта с другим по некоторому общему признаку. Метафора получается из сравнения элиминированием сравнительного слова «как». При осмыслении новых технологий в юридических терминах, метафоры, которые используются для их понимания, приобретают критически важное значение. Юрист постепенно начинает мыслить метафорами, привыкает к ним и рассматривает их как корректные сущностные понятия. И — ошибается!

Может ли содержательно быть интерпретирована «вина» ИИ как психическое отношение «комплекса технологических решений» к совершаемому им общественно опасному действию или бездействию и последствиям посредством интеллектуальных и волевых признаков? Уже на этом примере хорошо видна опасность метафор. Хочется надеяться, что Указом Президента РФ от 15.02.2024 года поставлена точка в бесплодных дискуссиях о «вине» и «ответственности» ИИ: «Не допускается делегирование системам искусственного интеллекта… ответственности за последствия принятия решений. Ответственность за все последствия работы систем ИИ всегда несет физическое или юридическое лицо, признаваемое субъектом ответственности в соответствии с законодательством РФ» (подп. „е“ п. 51.10)[4].

Суть «ошибки» ИИ также становится ясной из понятия «безошибочности». Синонимичным ему в Стратегии оказывается понятие «отказоустойчивость — способность технической системы сохранять работоспособность при отказе одной или нескольких ее составных частей» (подп. „с“ п. 5). И сразу становится ясной суть т. н. «ошибки» ИИ: неспособность системы управления с ИИ сохранять работоспособность при отказе даже одной ее составной части. То есть этот вопрос чисто технический.

Правовая база

Системы управления ядерным оружием могут рассматриваться как частный случай САСВ — «смертоносных автономных систем вооружений» («Lethal Autonomous Weapons Systems» — LAWS)[5]. В международном праве отсутствует специальный договор о применении САСВ в вооруженном конфликте. Тем более, нет соглашения об использовании смертоносных автономных систем управления ядерным оружием (далее — САСУЯО).

[1] О развитии искусственного интеллекта в Российской Федерации (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») : Указ Президента РФ от 10.10.2019 № 490 // ЮИС Легалакт. URL: https://legalacts.ru/doc/ukaz-prezidenta-rf-ot-10102019-n-490-o-razvitii/ (дата обращения: 15.04.2024).

[2] О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» и в Национальную стратегию, утвержденную этим Указом : Указ Президента Российской Федерации от 15 февраля 2024 г. № 124 // СПС «Гарант». URL: https://www.garant.ru/products/ipo/prime/doc/408459959/?ysclid=lv5i4dherm683992127 (дата обращения 15.04.2024).

[3] Там же.

[4] Там же.

[5] Views of the International Committee of the Red Cross (ICRC) on autonomous weapon system : 11 April 2016. [Par. 6].

Положения Женевских конвенций 1949 г. в области гуманитарного международного права, определяющие правила защиты людей при вооруженных конфликтах, солдат, раненых и военнопленных, а также гражданских лиц могут быть применены к рассматриваемым ситуациям лишь отчасти, равно как и оговорка Мартенса, которая запрещает использовать оружие, если это противоречит принципам гуманности, человечности и здравого смысла. Актуальность оговорки применительно к ядерному оружию отметил Международный суд ООН, назвав ее эффективным средством реагирования на стремительно меняющиеся военные технологии[1]. Таким образом, правовая база использования САСУЯО крайне бедна. Предлагаются разные решения. Некоторые авторы рекомендуют для САСВ разработать специальный международный договор [5, с. 27–28]. (Кстати, проект договора между СССР и США о взаимной компьютерной безопасности, охватывающий и обсуждаемую проблему, был опубликован еще в 1991 году [6, с. 145–153].) Представляется, что для САСУЯО это единственный вариант.

Итак, право вооруженных конфликтов и международное гуманитарное право дают лишь небольшое число опорных точек для решения проблемы. Вот одна из них. В международном гуманитарном праве есть группа связанных между собой терминов (все они начинаются с приставки *«re-»*, указывающей на повторяемость действия):

— *reprisal* — репрессалия, ответная мера;

— *reciprocity* — взаимность, соразмерность, взаимодействие;

— *revenge* — месть;

— *retaliation* — ответный удар, возмездие, воздаяние;

— *retortion* — реторсия, ответное действие (на недружественное действие).

Репрессалии — обдуманный противоправный акт в ответ на противоправный акт противника — неоднократно запрещены Дополнительным протоколом I (ДП I) с поправками 1977 года к Женевским конвенциям от 12 августа 1949 года, касающимся защиты гражданских жертв международной войны. О мести, слишком субъективной по своей природе, также не приходится говорить в правовом контексте. Реторсия — применение ограничительных мер одного государства в ответ на соответствующие действия другого государства по смыслу ближе к международно-правовым санкциям, и более того, применяются и в гражданском праве (ст. 1194 ГК РФ). Но возмездие, ответный удар (*retaliation*), с учетом взаимности и соразмерности (*reciprocity*) эмоционально менее нагруженные понятия означают, что этот исходный пункт «при определенных обстоятельствах может представлять собой меры и действия, являющиеся разумным и законным ответом на неразумные, хотя и законные эксцессы противоборствующей стороны» [7, с. 484].

В ДП I можно обнаружить еще две категории, весьма полезные для расширения топа: «опасные силы» (ст. 56) и «меры предосторожности при нападении» (ст. 57). В системном подходе Н. В. Щедрина «опасным силам» соответствует *источник опасности* — «свойство одной, чаще всего неустойчивой, системы (деятельности, ее объекта и субъекта), развитие и проявление которого поддается или не поддается контролю и с высокой вероятностью может произвести необратимые разрушительные изменения в этой или другой системе» [8, с. 28–29]. С технической точки зрения и российская система «Периметр», и американская система «Система экстренной связи для ракетного пуска» относятся к категории высоконадежных. Но именно высокий уровень недоверия каждой из сторон делает их в представлениях политиков и военных потенциально неустойчивыми. Для того чтобы снять это недоверие или, по крайней мере, минимизировать его и необходимо найти «топ безопасности».

«Мерам предосторожности при нападении» можно поставить в соответствие «меры безопасности» по Н. В. Щедрину. «Меры безопасности могут облекаться в нормативную оболочку «правил безопасности» и «санкций безопасности» [9, с. 89]. Меры безопасности, о которых можно договориться применительно двум указанным системам САСУЯО, и будут искомым топом безопасности. Н. В. Щедрин визуализировал меры безопасности в виде треугольника, разделенного на структурные нормативные элементы — правила (диспозиции) и санкции (ограничения) [1, с. 542]. «Правила безопасности — это совокупность обязанностей и запретов, которые субъект должен соблюдать, чтобы исключить или свести к минимуму вред, причиняемый источником повышенной опасности, либо предотвратить причинение ущерба объекту повышенной охраны любым источником опасности» [9, с. 86–87].

Таким образом, топ, который мы ищем (меры безопасности) в общей формулировке — это «реакция системы (комплекс реакций), необходимая и достаточная для того, чтобы ограничить вредное влияние источника опасности и сохранить возможность достижения целей, для которых эта система предназначена» [9, с. 84].

Н. В. Щедрин выделяет три вида санкций в структуре мер безопасности (рисунок 1): «санкция безопасности — «часть социальной нормы, в которой в качестве последствия общественно опасного поведения (деятельности) предусматривается ограничение возможностей продолжения такого поведения (деятельности)» [9, с. 88]; санкция наказания — «принудительное лишение определенных благ… путем угрозы или реального причинения лишения и страданий правонарушителю» [1, с. 543]; санкции восстановления — еще более сложное образование, которое «соединяет в себе ограничение, причинение страданий виновному и удовлетворение интересов жертвы» [9, с. 88–89]. Нетрудно видеть, что санкция безопасности соотносится с *«reciprocity»* (взаимность, взаимодействие) из ДП I,

[1] Консультативное заключение Международного Суда ООН от 8 июля 1996 г. по вопросу о законности угрозы ядерным оружием или его применения // Краткое изложение решений, консультативных заключений и постановлений Международного Суда. 1992–1996 годы. Нью-Йорк : ООН, 1998. URL: https://www.icj-cij.org/files/summaries/summaries-1992-1996-ru.pdf (дата обращения: 15.04.2024).

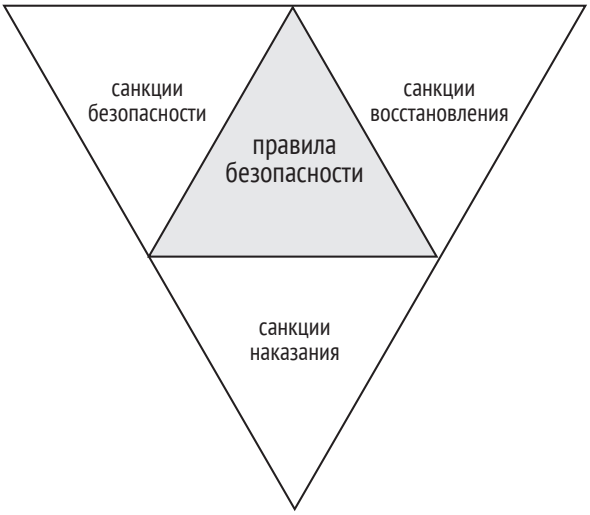

Рисунок 1 — Меры безопасности по Н. В. Щедрину [8, с. 30]

санкция наказания — с *«retaliation»* (возмездие), санкция восстановления — также с «retaliation» (возмездие), но не в материальном смысле, а другом значении — воздаяние, как выражение высшей справедливости и в этом смысле как «удовлетворение интересов жертвы».

Такова чрезвычайно узкая лингвистическая основа искомого топа.

Поможет расширить площадку взаимопонимания и снизить уровень недоверия экспериментальная юриспруденция.

Моделирование

В 1983–1985 гг. одновременно и независимо учеными СССР и США были созданы математические модели и получены прогнозы последствий ядерной войны, по одному из которых, климатическому (серьезному падению температур) проект получил название «ядерная зима» [10, с. 164–173]. Результаты были проверены моделированием процесса совместной исследовательской группой на американском суперкомпьютере с участием советских специалистов. Таким образом, было снято недоверие к советской модели. Именно это требуется сегодня в рассматриваемой ситуации недоверия к САСУЯО.

Первое приходящее на ум возражение: кто же предоставит для моделирования совместной работы совершенно секретные системы? В том-то и дело, что для предлагаемого эксперимента совершенно нет необходимости работать с данными конкретными системами. Речь идет об испытании и взаимной проверке принципов ответственного создания, управления, использования САСУЯО, что в конечном итоге выразится определенный уровень доверия к ним.

Итак, предлагается использовать решение уже привычное для юриспруденции — «песочницу», специально выделенную (изолированную) среду для безопасного исполнения компьютерных программ — с целью отработки на не стоящих на вооружении и разработанных специально для эксперимента моделях САСУЯО РФ и США. Тем самым секретность сохраняется, и по мере проведения эксперимента с помощью методов юридической топики вырабатываются нужные нормы, а площадка согласия постепенно расширяется.

Можно предположить, что найденные и сформулированные правила безопасности общего характера для САСУЯО будут выглядеть примерно следующим образом:

— «ИИ» не должен контролировать ядерное оружие без надзора и вмешательства человека;

— следует осторожно давать формулировки, каждый раз проверяя смысл используемых терминов, при разработке технического задания, создании алгоритмов, развертывании и использовании САСУЯО;

— целесообразно разрабатывать, создавать, развертывать и ставить на боевое дежурство САСУЯО только с использованием проверяемых методик, источников данных и процедур;

— необходимо минимизировать непреднамеренные искажения информации, ошибочные интерпретации и случайные помехи в контуре управления САСУЯО;

— требуется постоянное мониторирование качества выполнения критически важных функций САСУЯО и др.

Конечно, можно возразить, что подобные правила можно строить без всякого моделирования. Но в том-то и дело, что именно совместная работа и дает возможность находить формулировки, повышающие доверие. А самое главное, помимо такого рода общих формул обязательно появятся конкретные рекомендации по процедурам, методикам, мониторингу, найденные и согласованные именно в ходе совместного моделирования. Именно в этом суть топического метода.

Выводы

«Искусственный интеллект» сегодня в высокой степени мифологизирован, равно как и его «вина» за свои «ошибки». Как следствие, проблема безопасности систем с элементами «ИИ» приобрела мнимую составляющую, повышающую уровень недоверия к ним. А что касается САСУЯО, то именно для них уровень недоверия максимален. И это вполне понятно: нормы международного военного права в названной области отсутствуют, нормы международного гуманитарного права дают лишь подсказки. На площадке ООН в переговорах по безопасности в ИКТ-среде США и РФ исходят из разных принципов и их позиции непримиримы. Необходимо менять юридические подходы, и начать целесообразно с поиска «общего места» (топа) — цены «ошибки» САСУЯО. Путь к юридическому решению лежит через моделирование взаимодействия моделей двух автономных систем управления, специально создаваемых для совместного эксперимента в РФ и США. В процессе моделирования вырабатываются правила безопасности, которые впоследствии смогут составить предмет российско-американского соглашения.

Список источников

1. Уголовное право. Общая часть. Наказание : академический курс : в 10 томах. Т. 10. Иные уголовно-правовые меры. Меры безопасности и поощрение в уголовном праве / В. Н. Куфлева, С. В. Анощенкова, С. В. Полубинская [и др.]; рец.: И. М. Мацкевич, В. М. Хомич. Москва : Юрлитинформ, 2021. 759 с.

2. Соболева А. К. Топическая юриспруденция. Москва : Добросвет, 2001. 225 с.

3. Батурин Ю. М. Правонарушения в онлайновых вселенных и экспериментальная юриспруденция // Безопасность, конфликты и борьба с экстремизмом в информационном пространстве Интернета: правовые аспекты : материалы Международной научно-практической конференции (Москва, 1–2 июня 2017 г.). Москва : Водолей, 2017. С. 27–35.

4. Международная информационная безопасность: подходы России : доклад. Москва : МГИМО, 2021. 48 с.

5. Скуратова А. Ю., Королькова Е. Е. Смертоносные автономные системы вооружений: проблемы международно-правового регулирования // Российский юридический журнал. 2019. № 1 (124). С. 22–30.

6. Батурин Ю. М., Жодзишский А. М. Компьютерная преступность и компьютерная безопасность. Москва : Юридическая литература, 1991. 160 с.

7. Бест Дж. Война и право после 1945 г. Москва : ИРИСЭН ; Мысль, 2010. 676 с.

8. Щедрин Н. В. Концептуально-теоретические основы правового регулирования и применения мер безопасности // Криминология: вчера, сегодня, завтра. 2013. № 4 (31). С. 26–35.

9. Щедрин Н. В. Введение в правовую теорию мер безопасности. Красноярск : Красноярский государственный университет, 1999. 180 с.

10. Тарко А. М., Пархоменко В. П. Ядерная зима: история вопроса и прогнозы // Биосфера. 2011. Т. 3, № 2. С. 164–173.

КОНФЛИКТ ИНТЕРЕСОВ

Конфликт интересов отсутствует.

CONFLICT OF INTEREST

There is no conflict of interest.

Дата поступления статьи / Received: 22.04.2024.
Дата рецензирования статьи / Revised: 25.05.2024.
Дата принятия статьи к публикации / Accepted: 15.07.2024.