Научная статья
УДК 340.143:004.8
EDN: https://elibrary.ru/ekeudk
DOI: https://doi.org/10.21202/jdtl.2023.22

Этико-правовые модели взаимоотношений общества с технологией искусственного интеллекта

Дмитрий Валерьевич Бахтеев

Уральский государственный юридический университет имени В. Ф. Яковлева
г. Екатеринбург, Российская Федерация

Ключевые слова

ChatGPT,
искусственный интеллект,
машинное обучение,
модель,
общество,
право,
регулирование,
робот,
цифровые технологии,
этика

Аннотация

Цель: исследование современного состояния технологии искусственного интеллекта в формировании прогностических этико-правовых моделей взаимоотношений общества с рассматриваемой сквозной цифровой технологией.

Методы: основным методом исследования является моделирование. Помимо него, в работе использованы сравнительный, абстрактно-логический и исторический методы научного познания.

Результаты: сформулированы четыре этико-правовые модели взаимоотношений общества с технологией искусственного интеллекта: инструментальная (на основе использования человеком системы искусственного интеллекта), ксенофобная (на основе конкуренции человека и системы искусственного интеллекта), эмпатическая (на основе сочувствия и соадаптации человека и систем искусственного интеллекта), толерантная (на основе взаимоиспользования и сотрудничества между человеком и системами искусственного интеллекта). Приведены исторические и технические предпосылки формирования таких моделей. Описаны сценарии реакций законодателя на ситуации использования этой технологии, такие как необходимость точечного регулирования, отказа от регулирования либо же полномасштабного вмешательства в технологическую отрасль экономики. Произведено сравнение моделей по критериям условий реализации, достоинства, недостатков, характера отношений «человек – система искусственного интеллекта», возможных правовых последствий и необходимости регулирования отрасли либо отказа от такового.

Научная новизна: в работе приведена оценка существующих в научной литературе, публицистике мнений и подходов, проанализированы технические решения и проблемы, возникшие в недавнем прошлом и настоящем. Теоретические выводы подтверждаются ссылками на прикладные ситуации, имеющие общественную или правовую значимость. В работе использован междисциплинарный подход, объединяющий правовую, этическую и техническую составляющие, которые, по мнению автора, являются критериальными для любых современных социогуманитарных исследований технологий искусственного интеллекта.

Практическая значимость: феномен искусственного интеллекта связывают с четвертой промышленной революцией, соответственно, эта цифровая технология должна быть изучена многоаспектно и междисциплинарно. Выработанные в научной статье подходы могут быть использованы при дальнейших технических разработках интеллектуальных систем, совершенствования отраслевого законодательства (например, гражданского и трудового), а также при формировании и модификации этических кодексов в сфере разработки, внедрения и использования систем искусственного интеллекта в различных ситуациях.

Для цитирования

Бахтеев, Д. В. (2023). Этико-правовые модели взаимоотношений общества с технологией искусственного интеллекта. *Journal of Digital Technologies and Law*, *1*(2), 520–539. https://doi.org/10.21202/jdtl.2023.22

Содержание

Введение

Многочисленные достоинства систем искусственного интеллекта, среди которых быстрая обучаемость, возможность решать широкий круг задач, более высокая, чем у человека, эффективность вкупе со все большим проникновением их в различные сферы нашей жизни, заставляют задуматься над вопросами: способна ли (или будет ли способна) система искусственного интеллекта воспринимать себя как самостоятельную, независимую от разработчиков и пользователей личность, осознает ли искусственный интеллект свои преимущества перед людьми, как он будет оценивать свое положение и взаимодействие с человечеством, будет ли оно его устраивать и что он будет делать, если захочет его изменить. Эти вопросы находятся на пересечении предметности этики, права и технологии и потому подлежат разрешению междисциплинарными методами исследования (Kazim, 2021). В зависимости от ответов на эти вопросы и степени развития технологии возможны приведенные ниже модели/сценарии реакции общества и, как следствие, права на технологии искусственного интеллекта.

1. Инструментальная модель

Инструмент – продукт деятельности человека и средство изготовления других предметов, в том числе и инструментов. Карл Маркс и Мартин Хайдеггер в своих трудах подчеркивали необходимость различия между инструментом и машиной. Первый указывал на то, что зачастую смешиваются принципиально разные понятия: если инструмент требует непосредственного участия человека в процессе труда, то машина «заменяет рабочего, действующего одновременно только одним орудием, таким механизмом, который разом оперирует множеством одинаковых или однородных орудий и приводится в действие одной двигательной силой» (Маркс, 2001). Он же отмечает, что машину от инструмента отличает достаточная автономия (в большей степени ресурсная, машину в любом случае направляет и контролирует человек). Мартин Хайдеггер же, в свою очередь, с опорой на работы Гегеля, в качестве критериев отнесения объекта к машине называет самостоятельность, уверенность в себе и независимость (Хайдеггер, 1993). Для данного исследования это может пониматься следующим образом: искусственный интеллект в рамках инструментальной концепции на этапе разработки и тестирования выступает более как инструмент, а не как машина, поскольку его функционирование связано с человеческой деятельностью на трех уровнях: при разработке (в этом инструмент и машина схожи), при реализации деятельности, ранее присущей исключительно человеку, при контроле человеком результатов деятельности искусственного интеллекта. Такой подход иногда также называют прагматическим (Morley J. et al., 2021). «Машинность» искусственного интеллекта в данном случае является производной характеристикой от автономности, однако с точки зрения рассматриваемой модели информационная автономность систем искусственного интеллекта отрицается. Соответственно, в рамках данной модели искусственный интеллект рассматривается именно в качестве инструмента, в том числе для реализации нужд человечества (Watkins & Human, 2023).

Описанная в работах автора статьи теория автономности (информационной и ресурсной) (Бахтеев, 2021) вполне согласуется с тем, что системы искусственного интеллекта могут восприниматься и как машины (что, в свою очередь, подтверждается существованием терминов «машинное обучение» и «машинное зрение»), и как сущности, сопоставимые с познающими биологическими объектами. При этом следует учитывать, что если инструмент (в традиционном научном понимании) призван облегчить труд человека, то искусственный интеллект может заменить своей деятельностью труд человека. При этом некорректным будет сравнение искусственного интеллекта со станками времен промышленной революции. Такие станки лишили работы множество людей, однако при этом они создали новые рабочие места. В случае автоматизации вообще и использования систем искусственного интеллекта в частности мы уже наблюдаем ликвидацию определенных профессий, в первую очередь связанных с посредническими услугами: консультантов, диспетчеров, маркетологов и т. д. Обычные производственные роботы уже сейчас, без интеллектуальных модулей, в значительной степени оптимизировали конвейерное производство, что позволило при сокращении издержек на производство продукции многократно повысить количество и качество продукции, а также лишило работы огромное число неквалифицированных и низкоквалифицированных рабочих, фактически мы наблюдаем новую промышленную революцию. К 2020 г. в мире применялись более 3 млн промышленных роботов (к 2023 г. темпы роста, впрочем, снизились), происходит

интеграция интеллектуальных систем в разные области жизни, что очевидным образом положительно сказывается на экономике, но причиняет определенный ущерб обществу, в первую очередь в виде снижения занятости. Вариантом решения такой проблемы могут быть гарантии занятости, закрепленные законодательно, либо поддерживаемые государством и крупными компаниями программы переквалификации для людей, потерявших работу.

Развитие искусственного интеллекта позволяет решать ему все больший круг интеллектуальных задач, что создает предпосылки для дальнейшего сокращения рабочих мест во все большем количестве сфер, а также уничтожения целых профессий. «Реальный эффект сокращения фонда заработной платы (и числа рабочих мест. – Прим. Д. Б.) за счет внедрения роботов определяется количеством высвобождающихся при этом людей и величиной их заработной платы, а также стоимостью самих роботов, которая в свою очередь определяется сложностью конструкции и степенью интеллектуальности роботов» (Тимофеев, 1978).

В отличие от начального этапа роботизации распространение систем искусственного интеллекта может снизить уровень занятости не только для рабочих профессий. Так, согласно, прогнозам экспертов Оксфордского университета, в течение следующих 20 лет в США будет автоматизировано 47 %, а в Китае – 77 % рабочих мест. По мнению одного из ведущих специалистов в сфере вычислительной техники М. Варди, к 2045 г. примерно 50 % людей останутся без работы (Vardi, 2012). Можно возразить, что при этом увеличивается количество разработчиков искусственного интеллекта, однако увеличение числа программистов и других лиц, связанных с разработкой интеллектуальных систем, не идет в сравнение с уменьшением рабочих мест других профессий. Более того, стремление создать интеллектуальные системы, способные к саморепликации, также вполне очевидно, поэтому нельзя исключать сокращений и в профессиях, связанных с информационными технологиями: например, ChatGPT в своей четвертой версии способна создавать несложный, но компилируемый программный код. Так, в январе 2023 г. компания Alphabet объявила о сокращении 12 000 рабочих мест по всему миру, в том числе из-за внедрения различных интеллектуальных систем, способных заменить, в частности, маркетологов, копирайтеров и иллюстраторов[1]. Согласно модели, разработанной для анализа вероятности исчезновения отдельных профессий за счет использования интеллектуальных чат-систем, 100-процентное замещение человека на настоящем этапе развития технологии возможно для математиков, налоговых специалистов, финансовых аналитиков, писателей, копирайтеров, веб-дизайнеров, верстальщиков, секретарей государственных органов, новостных аналитиков (Eloundou T. et al., 2023).

К настоящему времени именно инструментальную модель можно считать единственной полноценно реализованной: в прикладной деятельности системы искусственного интеллекта выступают в качестве инструмента как средства повышения эффективности труда. Это, как и в случаях прежних технологических революций, накладывает на государство и общество функцию сохранения рабочих мест и регулирования интенсивности применения описываемой технологии.

[1] Pichar, S. (2023, January 20). *A difficult decision to set us up for the future*. https://blog.google/inside-google/message-ceo/january-update/

2. Ксенофобная модель

В связи с активизацией обсуждения вопросов разработки и использования систем искусственного интеллекта все чаще и громче слышатся голоса противников дальнейших исследований в данной области. При этом как минимум часть из них нельзя назвать обскурантами с иррациональным страхом перед технологическим прогрессом. Так, например, всемирно известный ученый, популяризатор науки С. Хокинг говорил: «…появление полноценного искусственного интеллекта может стать концом человеческой расы... Такой разум возьмет инициативу на себя и станет сам себя совершенствовать со все возрастающей скоростью. Возможности людей ограничены слишком медленной эволюцией, мы не сможем тягаться со скоростью машин и проиграем». Схожего мнения придерживается и известный американский инженер, IT предприниматель И. Маск, который заявил: «Полагаю, что искусственный интеллект рано или поздно прикончит всех нас... Facebook*, Google, Amazon, Apple – все они уже знают о вас очень многое. Искусственный интеллект, который будет создан в недрах этих корпораций, получит огромную власть над людьми. А концентрация власти в одних руках всегда порождает огромные риски». Отмечается также, что «распределение функций между системой искусственного интеллекта и человеком должно следовать принципу, ориентированному на человека, и оставлять всегда возможность для человеческого выбора. Это означает обеспечение человеческого контроля над рабочими процессами в системах искусственного интеллекта» (Семис-оол, 2019).

Данные факты детерминируют потребность тщательного исследования «ксенофобной» модели отношения к искусственному интеллекту. Следует указать, что она является развитием инструментальной модели по негативному сценарию, т. е. вследствие как прогресса в развитии и использовании систем искусственного интеллекта, так и реализации одного или нескольких рисков (в виде одномоментных негативных событий или же затяжных кризисов), описанных ранее.

Термин «ксенофобная» образован путем сочетания греческих слов ξένος («чужой») + φόβος («страх»). Тем самым буквально ксенофобия определяется как страх, нетерпимость к чужому, незнакомому[2].

Среди исследователей нет единства мнений об истоках происхождения ксенофобии. Ряд авторов отмечают, что она могла возникнуть как инструмент адаптации в процессе эволюции, который способствовал выживанию и передаче генов потомкам. Так, страх перед незнакомцами мог быть, среди прочего, основан на наблюдении, что чужаки могут быть разносчиками новых, а значит, очень опасных (из-за отсутствия нужных антител) для коренных жителей болезнетворных микроорганизмов.

Традиционно термин «ксенофобия» использовался для обозначения страха, неприязни к людям других рас, национальностей, культур и религий. Тем не менее, на наш взгляд, исследование процесса взаимодействия человечества с достижениями технологического прогресса дает основания заимствовать данный термин в том числе и для описания определенного типа отношения к ходу научно-технического прогресса и его плодам – технологиям, к числу которых очевидно относится и искусственный интеллект.

---

[2] Ожегов, С. И., Шведова, Н. Ю. (2016). *Толковый словарь русского языка* (с. 300). Москва: ООО «А ТЕМП».

Завершая изложение нашего подхода к ксенофобии, отметим важный аспект, который заключается в том, что в конечном итоге ксенофобия представляет собой специфическую разновидность страха. Страх, являясь одной из множества эмоций, по мнению Е. П. Ильина, выступает как «эмоциональное состояние, отражающее защитную биологическую реакцию человека или животного при переживании им мнимой или реальной опасности для их здоровья или благополучия» (Ильин, 2016). Далее Е. П. Ильин отмечает, что с биологической точки зрения страх – несомненно, полезное явление, в то время как для человека как социального существа страх нередко служит препятствием для достижения поставленных целей. В данной части работы будут исследованы основы потенциального критического недоверия общества к технологии искусственного интеллекта.

Сущность «ксенофобного подхода» к искусственному интеллекту заключается в рассмотрении его как реальной угрозы для человечества и его текущего положения в мире.

Обобщающий анализ критики технологии искусственного интеллекта позволяет выделить две основные формы страха (недоверия) людей к этой технологии – сущностную и инструментальную.

Сущностный страх связан с тем, что люди боятся не применения возможностей технологии искусственного интеллекта, а самого искусственного интеллекта как искусственного, но при этом вполне самостоятельного и автономного разума, способного существовать, учиться, мыслить и осознавать себя без участия человека. Появление такой «рукотворной мыслящей машины», которая способна мыслить не просто как человек, но и лучше человека, по существу, подрывает существовавшую на протяжении всей истории цивилизации монополию человека на мыслительную деятельность, которая и позволила ему занять доминантное положение среди всех других видов на планете. В этом смысле искусственный интеллект становится отдельным видом, который человечество не может воспринимать иначе как конкурентный. При этом страх вызывает именно вышедший из-под контроля искусственный интеллект, т. е. ситуация обретения искусственным интеллектом самосознания в результате программного сбоя или умышленных действий разработчика. Фактически такую ситуацию, проводя аналогию с биологическими процессами, следует считать мутацией, однако сомнительно, что процессы разработки технологических продуктов имеют настолько большую связь с эволюционными механизмами. Именно поэтому сценарий «агрессивного» искусственного интеллекта представляется крайне нереалистичным.

Инструментальный страх, в свою очередь, отражает боязнь вытеснения человека системами искусственного интеллекта в трудовой сфере, что было описано в инструментальной модели (см. выше).

После начала систематического исследования искусственного интеллекта в 40–50-е гг. XX в. не прошло и ста лет, как уже созданы системы, превосходящие человека в отдельных областях интеллектуальной деятельности. Возможности современных компьютеров пока не позволяют произвести полное моделирование сознания человека или всего окружающего мира, однако с абстракциями искусственный интеллект справляется. Игры – это вполне корректная абстракция и, что в данном случае крайне важно, результаты участия в игре могут быть точно оценены. Так, уже с начала 2000-х гг. сильнейшие игроки в шахматы в мире ничего не могут противопоставить компьютеру, а по словам Г. Каспарова, все профессиональные шахматисты

тренируются, играя против компьютерных шахматных программ, поскольку соперник-человек не может предоставить достаточной глубины вариантов ходов. В 2015 г. компьютерная программа впервые победила человека в игре го – одной из самых сложных игр с открытой информацией[3], что ранее считалось невозможным. Искусственные нейросети способны обыгрывать профессиональных игроков в компьютерные игры в киберспортивных дисциплинах[4], что также считалось ранее исключительной прерогативой человека.

Содержание ксенофобного подхода заключается в том, что искусственный интеллект может использоваться отдельными лицами, организациями и целыми государствами как средство достижения своих недобросовестных целей.

Типичным примером данного страха является скандал, возникший вскоре после президентских выборов в США, связанный с организацией Cambridge Analytica. Согласно информации из ряда источников, данная частная организация, используя новейшие методы сбора и анализа данных в социальной сети Facebook*, получила огромный массив данных, в том числе и личных, для разработки специальной политической рекламы, которая, по мнению ряда экспертов, в значительной степени способствовала избранию на президентскую должность нынешнего президента США. Более того, данная организация обвиняется в причастности к вмешательству в результаты более 200 выборов по всему миру. Бывший сотрудник Cambridge Analytica, К. Уайли, отмечал: «Мы использовали несовершенство программного обеспечения Facebook* для сбора миллионов пользовательских профилей и построения моделей, которые позволяли нам узнавать о людях и применять эти знания для активации их внутренних демонов»[5].

Данное событие демонстрирует, что уже сейчас возможности систем искусственного интеллекта применяются для сбора персональных данных, манипуляции общественным мнением. В будущем искусственный интеллект может использоваться для манипуляции огромными объемами информации, формируя картину мира для населения целых государств, что создает реальные угрозы для демократических институтов, свободы слова и распространения информации. Этим могут воспользоваться государства для навязывания определенной позиции своему населению, а также населению других стран; представители различных корпораций для искусственного формирования спроса на те или иные товары и услуги. Наконец, этим могут воспользоваться представители преступного мира для сбора конфиденциальной информации граждан и организаций, которая впоследствии может быть продана на черном рынке или использована для шантажа или мошеннических действий.

---

3 Так, в шахматах на каждый шаг игры существует около 20 возможных реакций-ответов, в го их около 200.

4 Statt, N. (2019, April 13). OpenAI's Dota 2 AI steamrolls world champion e-sports team with back-to-back victories. *The Verge*. https://www.theverge.com/2019/4/13/18309459/openai-five-dota-2-finals-ai-bot-competition-og-e-sports-the-international-champion

5 Черешнев, Е. (20 марта 2018). Беззащитные данные: как Facebook оказалась в центре самого большого скандала в истории. *Forbes*. https://www.forbes.ru/tehnologii/358883-bezzashchitnye-dannye-kak-facebook-okazalas-v-centre-samogo-bolshogo-skandala-v

Иной пласт данной проблемы заключается в применении искусственного интеллекта в военных действиях. Условный боевой робот, снабженный искусственным интеллектом, или даже целая армия таких роботов представляет собой эффективную замену действующим регулярным войскам. Он способен выполнять сложные интеллектуальные задачи, может действовать даже в самых неблагоприятных условиях, не требует сна и отдыха, а его уничтожение мало волнует общественное мнение в стране, ведущей войну (по крайней мере, до тех пор, пока такая война не становится слишком дорогой с экономической точки зрения). При этом искусственный интеллект способен решать ранее исключительно человеческие задачи с внечеловеческой, машинной рациональностью. Для системы искусственного интеллекта при неверном ее формировании не возникнет проблемы использования незаконных средств ведения войны, убийства гражданского населения и т. п.

Иной аспект ксенофобного подхода связан с явлениями, описанными в части инструментального подхода. Так, голливудская гильдия сценаристов, а также тысячи художников и иллюстраторов по всей планете считают результаты работы системы искусственного интеллекта априорно плагиатом, поскольку они не представляют творчество как таковое, а лишь являются перемешиванием уже раскрытых смыслов. Отметим, однако, что значительная часть человеческого творчества формируется по аналогичной модели.

Таким образом, ксенофобный подход к оценке искусственного интеллекта нельзя назвать однозначно необоснованным. Существуют вполне резонные основания опасаться неконтролируемого развития технологий искусственного интеллекта и все большей интеграции их в разные аспекты жизни человека. Данный подход, предполагающий либо жесткий контроль над исследованиями в данной области, либо, в более радикальной форме, полный отказ от таковых, очевидно, не лишен недостатков. К их числу относятся: торможение научно-технического прогресса, невозможность оптимизации практической деятельности людей посредством использования систем искусственного интеллекта, разрыв между научными достижениями и их интеграцией в практику и, как следствие, снижение авторитета компьютерной науки в глазах общественности. Альтернативой цифровым технологиям, флагманом которых является искусственный интеллект, обычно считаются биотехнологии, поэтому отказ от развития искусственного интеллекта, разочарование в этой технологии могут привести к развитию медицины, физиологии, генетики и т. д. Тем не менее следует, на наш взгляд, не забывать о достоинствах данного подхода, который предполагает более взвешенную оценку технологии и ее прикладного значения; разработку инструментария для прогнозирования и оценки рисков дальнейшего изучения и использования искусственного интеллекта, который с поправками может быть использован и в других областях знания; стимулирование развития иных наук, связанных с развитием человека и его потенциала; придание нового импульса осмыслению человека и его места в мире.

3. Эмпатическая модель

Согласно этой модели, общество, воспринимая положительно домашних и иных социально-бытовых интеллектуальных роботов и программных помощников, благоприятно воспринимает идею распространения технологии и не исключает возможности наделения систем искусственного интеллекта свойствами субъекта права (в ограниченном смысле).

В основе данной модели лежит продвинутое и расширенное ощущение гуманизма и ответственности человека не только за себя, но и за тех, кто рядом. Интеллектуальные и автономные системы, согласно данной модели, перестают считаться инструментом или конкурентом человеку, а рассматриваются как компаньоны, но в ограниченном смысле, на уровне животных-компаньонов. Именно этические и правовые нормы, регулирующие отношение к животным, являются фундаментом реализации этой модели. Фактически данная модель является переходной между инструментальной и толерантной и не может рассматриваться как нечто продолжительное.

Приведем примеры, подтверждающие, что частично данная модель уже реализуется в обществе.

В работе К. Дарлинг и С. Хоерт описывается следующий эксперимент. Отдельным шести группам по восемь человек выдаются игрушки в форме динозавра, размером с небольшую кошку. Участникам предлагалось взаимодействовать с ними. Затем каждой группе была дана команда «задушить», «разбить голову» или иным способом «убить» игрушки, которая встретила жесткое сопротивление: участники не просто отказывались «убить» «своих» динозавров, но и пытались защищать их от других людей и испытывали серьезный дискомфорт, если видели, как «умирает» динозавр[6]. При этом из 48 игрушек только одна была «убита».

Другим примером может служить отключение серверов, обслуживающих роботов-игрушек Jibo, которое многие пользователи восприняли как смерть своего компаньона и отреагировали крайне эмоционально.

Аналогичным образом люди часто реагируют на проблемы и смерть книжных, кинотеатральных и игровых персонажей, несмотря на то, что существуют они исключительно виртуально.

В психологии существует позиция, что людям нравится общаться с чат-ботами (такими как ChatGPT) для обсуждения своих психологических проблем, поскольку интеллектуальная система редко способна на осуждение, свойственное живому человеку[7]. Надо полагать, в будущем это явление будет встречаться еще чаще.

Приведенные ситуации, несмотря на то, что они являются частными случаями и не могут отражать в целом общественное отношение к интеллектуальным системам, демонстрируют, что в определенных случаях человек и группы людей могут относиться к явно неодушевленным (и, стоит признать, даже не вполне интеллектуальным) системам как к любимым питомцам. Отметим также, что чувство эмпатии напрямую зависит от внешнего вида (внешности?) киберфизической системы и используемой лексики. Так, степень эмпатии и доверия к киберфизической системе антропоморфного фенотипа может также зависеть от черт «лица». М. Б. Матур и Д. Б. Райхлинг обнаружили, что надежность робота варьируется в зависимости от сходства его лица с человеком, не увеличивается линейно с человеческим подобием, но падает, когда агент очень реалистичен, но еще не полностью похож на человека (Mathur & Reichling, 2016). Этот феномен, изначально описанный в 1978 г.

---

6 См. Darling, K., Hauert, S. (2013, March 8). Giving rights to robots. *Robohub.org*. RobotsPodcast № 125. https://robohub.org/robots-giving-rights-to-robots/

7 Эксперт: Люди используют нейросеть в качестве психолога из-за страха осуждения со стороны живого человека. (2023, 23 марта). *Агентство городских новостей «Москва»*. https://www.mskagency.ru/materials/3286743

японским ученым М. Мори, носит название «странная долина», или «зловещая долина» (uncanny valley): самые необычные человекоподобные роботы неожиданно оказались неприятными людьми из-за неуловимых несоответствий во внешности и поведении, вызывающих чувство дискомфорта и страха (Mori, 2012). Таким образом, существует вероятность, что с развитием робототехники эмпатическая модель может сдвинуться не в сторону толерантной, а по направлению ко ксенофобной. Так, риски именно психологического характера были отражены в Своде этических правил для роботов, разработанном Британским институтом стандартов: пользователю интеллектуальной системы не должно быть неуютно, он не должен испытывать тревогу и стресс (Winfield, 2019).

Уточнение данной модели требует также раскрытия феномена взаимообучения человека и машины. Взаимодействуя с интеллектуальными системами, человек меняется, при этом изменения затрагивают не только сферу технологических навыков, но также физиологию и морально-этическую сферу. Так, исследование группы швейцарских ученых привело к экспериментально проверенному выводу о том, что «повторяющиеся движения по гладкой поверхности сенсорного экрана изменяют сенсорные реакции и, как следствие, представления мозга о последствиях прикосновений» (Balerna & Ghosh, 2018): мозг современного человека при прикосновении пальцев рук к поверхности с интенсивностью, аналогичной управлению смартфоном, ожидает изменения «картинки» перед глазами.

К другим особенностям, имеющим значение для рассмотрения данной модели, следует отнести ухудшение памяти и концентрации внимания, обусловленное возможностью быстрого обращения к необходимой информации с помощью, к примеру, смартфона или голосового помощника, а также улучшение визуальных навыков, позволяющих быстрее осуществлять более качественное комплексное восприятие сложных визуальных объектов. Не следует в целом считать, что интеграция интеллектуальных систем в общество приводит к деградации последнего. Эффект Флинна демонстрирует, что средний интеллектуальный уровень каждого нового поколения возрастает, т. е. текущий средний коэффициент интеллекта соответствует высокому уровню интеллекта предшествующего поколения (Flynn, 2009). При этом современные исследования показывают, что с распространением цифровых технологий действие эффекта Флинна снизилось или вообще исчезло (Teasdale, 2005). Однако данное исследование не может считаться достоверным, так как его объектом являлся интеллект призывников, т. е. выводы могут объясняться социальными причинами, а не реальным падением интеллектуального уровня. Представляется, что имеет место не снижение интеллекта, а его перепрофилирование (Букатов, 2018), к примеру, современный студент помнит меньше своих предшественников из XX в., однако владеет гораздо большим набором техник для поиска и анализа информации. Соответственно, наблюдается постепенная замена содержательных знаний навыками работы с информацией, «следует учитывать биологическую коадаптацию и коэволюцию органов чувств человека, следует учитывать расширение диапазона наших восприятий, которое обеспечивается благодаря достижениям техники» (Огурцов, 2006). Однако нельзя исключать факторы притупления внимания, снижения понимаемой ответственности и профессионализма лиц, принимающих решения, чьим «советником» является система искусственного интеллекта. Таким образом, возможна ситуация «перекладывания» ответственности за ошибку/противоправное деяние на систему искусственного интеллекта.

Помимо физиологического и интеллектуального аспекта, эмпатическая модель включает в себя и возможные изменения в эмоциональной сфере, так, М. Шойц указывает: «Социальные роботы, которые устанавливают с людьми эмоциональный контакт и, как следствие, последние глубоко доверяют роботам, что в свою очередь может быть использовано для манипулирования людьми ранее невозможными способами. Например, компания может использовать уникальные отношения робота со своим владельцем, чтобы робот убедил владельца приобрести продукты, которые компания хочет продвигать. Обратите внимание на отличие от человеческих отношений, где при нормальных обстоятельствах социальные эмоциональные механизмы, такие как эмпатия и вина, предотвратят эскалацию таких сценариев» (Scheutz, 2009).

Как известно, во многих странах предусмотрена уголовная ответственность за проявленную по отношению к животным жестокость и тем более лишение жизни животного без оснований на это, так, к примеру, Европейская конвенция о защите домашних животных указывает на недопустимость причинения без необходимости боли и страданий[8]. Продолжая сравнение систем искусственного интеллекта с домашними животными, не будет ли существенное перепрограммирование такой системы причинять ей боль, аналогично тому, как боль животному причиняет косметическое хирургическое вмешательство, также запрещаемое указанной Конвенцией?

Соответственно, в рамках рассматриваемой модели такой подход переносится на системы искусственного интеллекта и с большими оговорками действует в настоящее время в части общества. Однако для полноценной его реализации необходим хотя бы условный и воспринятый обществом ответ на вопрос о том, может ли система искусственного интеллекта испытывать боль и страдания.

4. Толерантная модель

Поступательное развитие научно-технического прогресса при сохранении интереса к совершенствованию рассматриваемых технологий может привести к указанной ранее ситуации – появлению «сильного» искусственного интеллекта или как минимум широкому распространению интеллектуальных систем-помощников. В первом случае ограничения технического характера будут нивелированы, системы искусственного интеллекта получат достаточную автономию, единственным рамочным ограничением которой смогут послужить правовые нормы и производные от них технические ограничения. Такое условное равенство (или как минимум соотносимость) человечества и искусственного интеллекта может повлечь за собой как положительные, так и отрицательные явления.

В рамках данной модели искусственный интеллект выступает «партнером» человечества, обеспечивает функцию сдерживающего механизма, не допускает эскалации конфликтов, осуществляет общую и частную превенцию правонарушений. Негативные сценарии, описанные ранее, не реализованы, поскольку благополучное

[8] *European Convention for the Protection of Pet Animals* (1987, November 13). Council of Europe: official website. https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168007a67d

существование как человечества, так и искусственного интеллекта взаимозависимы, и обе указанные сущности обеспечивают и собственную стабильность, и развитие (в случае человечества в первую очередь социальное, для систем искусственного интеллекта – техническое) за счет сотрудничества, но не конкуренции. «Вовлеченность в рыночные отношения требует взаимного учета интересов и прав… Трудно назвать какой-либо иной, помимо взаимоиспользования, принцип, посредством которого равенство и справедливость утверждались в человеческих отношениях столь естественно и спонтанно» (Апресян, 1995). Отношения общества и систем искусственного интеллекта вряд ли можно назвать в полном смысле человеческими, однако ожидать от них взаимности, основанной на прямой и обратной полезности, вполне оправданно. Так, в Национальном стандарте «Требования по безопасности для промышленных роботов» используются следующие формулировки: «взаимодействующий с человеком робот (collaborative robot): Робот, сконструированный для непосредственного взаимодействия с человеком в рамках определенного совместного рабочего пространства; совместное рабочее пространство (collaborative workspace): Рабочее пространство, находящееся внутри защищенного пространства, где робот и человек могут выполнять работы одновременно в процессе производства»[9]. Очевидно, что робот в данных определениях является не инструментом, а субъектом взаимодействия, коллаборации, совместной работы, что, однако, пока выглядит преждевременным.

Так, успехи систем искусственного интеллекта в научной и творческой деятельности могут привести к росту доходов, которые направляются на дальнейшее развитие системы искусственного интеллекта. Проекты на основе искусственного интеллекта становятся не просто самоокупаемыми, но и сверхдоходными. Следует, однако, учитывать, что любой экономический рост ограничен как в объемах, так и во времени.

Развитие технологии может повлечь качественное изменение жизни: искусственный интеллект (к примеру, в качестве автора произведений искусства или изобретения) представляет свои интересы в суде, создает конкуренцию представителям различных профессий, что в свою очередь формирует ситуацию конкуренции и стимул для развития общества и человека.

Толерантная модель очевидным образом остается единственной доступной из приведенного списка при создании сильного искусственного интеллекта, однако она также может быть реализована в ближайшем будущем, даже в условиях отсутствия последнего: при возрастании числа ситуаций эффективного безрискового практического использования систем искусственного интеллекта возрастет доверие к ним на корпоративном и государственном уровнях. К примеру, канадская страховая компания Kanetrix.ca использует такие системы в процессе выбора пользователем приобретаемого страхового продукта. Учитывая, что для этого используются искусственные нейронные сети, прозрачности и объяснимости ожидать было бы странно, однако эти характеристики и не требовались: из

[9] *ГОСТ Р 60.1.2.2-2016/ИСО 10218-2:2011 Роботы и робототехнические устройства. Требования по безопасности для промышленных роботов. Часть 2: Робототехнические системы и их интеграция.* (2016). Москва: Стандартинформ.

сопоставления с человеком система искусственного интеллекта в данном случае убедительно вышла победителем[10].

Подобные состояния могут стать реальностью лишь при условии полного разрешения вопроса границ ответственности систем искусственного интеллекта, определения критериев наличия сознания и волевого компонента в принятии решения, без которых не может произойти наделение систем искусственного интеллекта признаками субъекта права.

Недостатки данной модели находятся в плоскости как публичного, так и частного права: признание обществом систем искусственного интеллекта как равного человеку субъекта невозможно без ответа на вопросы о самой сути и критериях наложения на такую систему ответственности за ее действия, т. е. отнесение ее к объекту или субъекту права. Так, В. А. Лаптев считает, что последствия действий и решений искусственного интеллекта могут рассматриваться как обстоятельство непреодолимой силы, т. е. исключающее саму постановку вопроса об ответственности, либо же введение обязательного страхования гражданской ответственности создателя системы искусственного интеллекта за вред, причиненный последними третьим лицам (Лаптев, 2017).

Если системы искусственного интеллекта (или киберфизические системы) достигнут определенного уровня когнитивных способностей, т. е. если они будут обладать очевидной моральной значимостью, такой как разум или чувствительность, тогда они, вероятно, будут претендовать на признание их морального статуса и должны иметь права, т. е. некоторую долю привилегий, претензий, полномочий или иммунитетов (Gunkel, 2018). Данная ситуация возможна лишь при значительном качественном прогрессе технологии. Очевидно, что для описания данной модели используется слишком много слов «если» и «вероятно». Это выражает степень неуверенности в возможности развития технологии искусственного интеллекта до таких пределов, однако исключать такую возможность все же нельзя. Исходя из темпов развития технологии, эксперты прогнозируют, что при самом худшем стечении обстоятельств полноценный искусственный интеллект, способный сравниться с человеком, будет разработан к концу XXI в.

Выводы

В обобщенном виде соотношение описанных моделей может быть представлено в табличной форме (табл.).

Отметим также, что данные модели не отражают последовательность развития общественных отношений в контексте технологий машинного обучения, они могут реализовываться параллельно друг другу в разных регионах, отраслях экономики, судопроизводства и т. д.

[10] McWaters, R. J. et al. *Navigating Uncharted Waters. A roadmap to responsible innovation with AI in financial services.* World Economic Forum. http://www3.weforum.org/docs/WEF_Navigating_Uncharted_Waters_Report.pdf

Соотношение моделей взаимодействия общества и права с технологией искусственного интеллекта

Критерий сравнения	Инструментальная модель	Ксенофобная модель	Эмпатическая модель	Толерантная модель
Условие реализации	Реализована в настоящее время	При проявлении существенных кризисов	Развитие эмпатических настроений в обществе, прогресс в робототехнике и интеллектуальных системах-ассистентах	Достижение технологической сингулярности, появление «сильного (общего)» искусственного интеллекта
Достоинства	Низкий уровень социальных и правовых рисков от использования систем искусственного интеллекта при обеспечении промышленного и интеллектуального прогресса, возможность сохранения текущих подходов к регулированию технологий	Развитие медицины, физиологии, генетики	Развитие общественной морали и гуманизма (в широком смысле)	Развитие как технологий, так и права и других социогуманитарных областей знания
Недостатки	Застой в социогуманитарных науках, отказ от концепции «сильного (общего)» искусственного интеллекта, антропоцентризм	Торможение научно-технического прогресса в области цифровых и компьютерных технологий	Уменьшение рациональности в угоду эмоциональности, трансформации эмоциональной сферы, массовые проблемы с запоминанием и вниманием	Снижение требований к эффективности систем искусственного интеллекта. Возможность использования системы искусственного интеллекта, обладающей правосубъектностью как «подставного лица», требование пересмотра большого количества правовых и иных социальных норм
Характер отношений «человек – искусственный интеллект»	Использование	Конкуренция	Сочувствие, соадаптация	Взаимоиспользование, сотрудничество
Правовые последствия	За негативные решения и действия системы искусственного интеллекта отвечает человек (оператор или разработчик)	Введение разрешительного регулирования отрасли	Повышенная правовая защита интеллектуальных систем без наделения их правосубъектностью	Наделение систем искусственного интеллекта правосубъектностью

Подводя итог, подчеркнем необходимость развития изучения потенциала систем искусственного интеллекта, в том числе их свойств при интеграции, и распространения в обществе, а также потенциальных рисков, неизбежных при этих процессах. Указанные модели отражают грани реальности, как уже существующей, так и потенциальной. Моделирование подобных ситуаций следует осуществлять дифференцированно, чему могут поспособствовать приведенные этико-правовые модели.

* Организация признана экстремистской, ее деятельность запрещена на территории Российской Федерации

Список литературы

Апресян, Р. Г. (1995). Нормативные модели моральной рациональности. В кн. *Мораль и рациональность* (с. 94–118). Москва: Институт философии РАН.

Бахтеев, Д. В. (2021). *Искусственный интеллект: этико-правовой подход*: монография. Москва: Проспект.

Букатов, В. М. (2018). Клиповые изменения в восприятии, понимании и мышлении современных школьников – досадное новообразование «постиндустриального уклада» или долгожданная реанимация психического естества? *Актуальные проблемы психологического знания*, *4*(49), 5–19.

Ильин, Е. П. (2016). *Эмоции и чувства*. (2-е изд.). Санкт-Петербург: Питер.

Лаптев, В. А. (2017). Ответственность «будущего»: правовое существо и вопрос оценки доказательств. *Гражданское право*, *3*, 32–35.

Маркс, К. (2001). *Капитал* (Т. 1). Москва: АСТ.

Огурцов, А. П. (2006). Возможности и трудности в моделировании интеллекта. В кн. Д. И. Дубровский, В. А. Лекторский (ред.) *Искусственный интеллект: междисциплинарный подход* (с. 32–48). Москва: ИИнтеЛЛ.

Семис-оол, И. С. (2019). «Заслуживающий доверия» искусственный интеллект. В сб. Д. В. Бахтеев, *Технологии XXI века в юриспруденции: мат-лы Всерос. науч.-практ. конф. (Екатеринбург, 24 мая 2019 года)* (с. 145–149). Екатеринбург: Уральский государственный юридический университет.

Тимофеев, А. В. (1978). *Роботы и искусственный интеллект*. Москва: Главная редакция физико-математической литературы издательства «Наука».

Хайдеггер, М. (1993). Вопрос о технике. В сб. *Время и бытие: статьи и выступления*. Москва: Республика.

Balerna, M., & Ghosh, A. (2018). The details of past actions on a smartphone touchscreen are reflected by intrinsic sensorimotor dynamics. *Digital Med*, *1*, Article 4. https://doi.org/10.1038/s41746-017-0011-3

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023, Marth 17). GPTs are GPTs: An *Early Look at the Labor Market Impact Potential of Large Language Models*.

Flynn, J. R. (2009). *What Is Intelligence: Beyond the Flynn Effect*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511605253

Gunkel, D. J. (2018). *Robot rights*. Cambridge, MA: MIT Press.

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns*, *3*(9). https://doi.org/10.1016/j.patter.2021.100314

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition, 146,* 22–32. https://doi.org/10.1016/j.cognition.2015.09.008

Mori, M. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/mra.2012.2192811

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mokander, J., & Floridi, L. (2021). Ethics as a service: a pragmatic operationalisation of AI Ethics. *Minds and Machines*, *31*, https://doi.org/10.1007/s11023-021-09563-w

Scheutz, M. (2009). The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. *Workshop on Roboethics at ICRA*.

Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: *The Flynn Effect in reverse. Personality and Individual Differences*, *39*(4), 837–843. https://doi.org/10.1016/j.paid.2005.01.029

Vardi, M. (2012). Artificial Intelligence: Past and Future. *Communications of the ACM*, *55*, 5. https://doi.org/10.1145/2063176.2063177

Watkins, R., & Human, S. (2023). Needs-aware artificial intelligence: AI that 'serves [human] needs. *AI Ethics*, *3*. https://doi.org/10.1007/s43681-022-00181-5

Winfield, A. (2019). Ethical standards in robotics and AI. *Nature Electronics*, *2*, 46–48. https://doi.org/10.1038/s41928-019-0213-6

Сведения об авторе

Бахтеев Дмитрий Валерьевич – доктор юридических наук, доцент, доцент кафедры криминалистики, Уральский государственный юридический университет имени В. Ф. Яковлева, руководитель группы проектов CrimLib.info
Адрес: 620137, Российская Федерация, г. Екатеринбург, ул. Комсомольская, 21
E-mail: ae@crimlib.info
ORCID ID: https://orcid.org/0000-0002-0869-601X
ScopusAuthorID:https://www.scopus.com/authid/detail.uri?authorId=57208909117
Web of Science Researcher ID:
https://www.webofscience.com/wos/author/record/ABA-1494-2020
Google Scholar ID: https://scholar.google.ru/citations?user=h0zOOdcAAAAJ
РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=762765

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.07.49 / Планирование и прогнозирование в праве
Специальность ВАК: 5.1.1 / Теоретико-исторические правовые науки

История статьи

Дата поступления – 25 марта 2023 г.
Дата одобрения после рецензирования – 24 апреля 2023 г.
Дата принятия к опубликованию – 16 июня 2023 г.
Дата онлайн-размещения – 20 июня 2023 г.

Research article
DOI: https://doi.org/10.21202/jdtl.2023.22

Ethical-Legal Models of the Society Interactions with the Artificial Intelligence Technology

Dmitriy V. Bakhteev

Ural State Law University named after V. F. Yakovlev
Ekaterinburg, Russian Federation

Keywords

Artificial Intelligence,
ChatGPT,
digital technologies,
ethics,
law,
machine learning,
model,
regulation,
robot,
society

Abstract

Objective: to explore the modern condition of the artificial intelligence technology in forming prognostic ethical-legal models of the society interactions with the end-to-end technology under study.

Methods: the key research method is modeling. Besides, comparative, abstract-logic and historical methods of scientific cognition were applied.

Results: four ethical-legal models of the society interactions with the artificial intelligence technology were formulated: the tool (based on using an artificial intelligence system by a human), the xenophobia (based on competition between a human and an artificial intelligence system), the empathy (based on empathy and co-adaptation of a human and an artificial intelligence system), and the tolerance (based on mutual exploitation and cooperation between a human and artificial intelligence systems) models. Historical and technical prerequisites for such models formation are presented. Scenarios of the legislator reaction on using this technology are described, such as the need for selective regulation, rejection of regulation, or a full-scale intervention into the technological economy sector. The models are compared by the criteria of implementation conditions, advantages, disadvantages, character of "human – artificial intelligence system" relations, probable legal effects and the need for regulation or rejection of regulation in the sector.

Scientific novelty: the work provides assessment of the existing opinions and approaches, published in the scientific literature and mass media, analyzes the technical solutions and problems occurring in the recent past and present. Theoretical conclusions are confirmed by references to applied

situations of public or legal significance. The work uses interdisciplinary approach, combining legal, ethical and technical constituents, which, in the author's opinion, are criteria for any modern socio-humanitarian researches of the artificial intelligence technologies.

Practical significance: the artificial intelligence phenomenon is associated with the fourth industrial revolution; hence, this digital technology must be researched in a multi-aspectual and interdisciplinary way. The approaches elaborated in the article can be used for further technical developments of intellectual systems, improvements of branch legislation (for example, civil and labor), and for forming and modifying ethical codes in the sphere of development, introduction and use of artificial intelligence systems in various situations.

For citation

Bakhteev, D. V. (2023). Ethical-legal Models of the Society Interactions with the Artificial Intelligence Technology. *Journal of Digital Technologies and Law, 1*(2), 520–539 https://doi.org/10.21202/jdtl.2023.22

References

Apresyan, R. G. (1995). Normative models of moral rationality. In *Morals and rationality* (pp. 94–118). Moscow: Institut filosofii RAN. (In Russ.).

Bakhteev, D. V. (2021). *Artificial intelligence: ethical-legal approach*. Moscow: Prospekt. (In Russ.).

Balerna, M., & Ghosh, A. (2018). The details of past actions on a smartphone touchscreen are reflected by intrinsic sensorimotor dynamics. *Digital Med*, *1*, Article 4. https://doi.org/10.1038/s41746-017-0011-3

Bukatov, V. M. (2018). Clip changes in the perception, understanding and thinking of modern schoolchildren – negative neoplasm of postindustrial way or long-awaited resuscitation of the psychic nature? *Actual Problems of Psychological Knowledge*, *4*(49), 5–19. (In Russ.).

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023, March 17). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*.

Flynn, J. R. (2009). *What Is Intelligence: Beyond the Flynn Effect*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511605253

Gunkel, D. J. (2018). *Robot rights*. Cambridge, MA: MIT Press.

Heidegger, M. (1993). The Question Concerning Technology. In *Time and being: articles and speeches*. Moscow: Respublika. (In Russ.).

Ilyin, E. P. (2016). *Emotions and feelings*. (2d ed.). Saint Petersburg: Piter. (In Russ.).

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns*, *3*(9). https://doi.org/10.1016/j.patter.2021.100314

Laptev, V. A. (2017). Responsibility of the "future": legal essence and evidence evaluation issue. *Civil Law*, *3*, 32–35. (In Russ.).

Marx, K. (2001). *Capital* (Vol. 1). Moscow: AST. (In Russ.).

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition, 146,* 22–32. https://doi.org/10.1016/j.cognition.2015.09.008

Mori, M. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/mra.2012.2192811

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mokander, J., & Floridi, L. (2021). Ethics as a service: a pragmatic operationalisation of AI Ethics. *Minds and Machines*, *31*. https://doi.org/10.1007/s11023-021-09563-w

Ogurtsov, A. P. (2006). Opportunities and difficulties in modeling intelligence. In D. I. Dubrovskii, & V. A. Lektorskii (Eds.), *Artificial intelligence: interdisciplinary approach* (pp. 32–48). Moscow: IIntELL. (In Russ.).

Scheutz, M. (2009). The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. *Workshop on Roboethics at ICRA*.

Semis-ool, I. S. (2019). "Trustworthy" artificial intelligence. In D. V. Bakhteev (Ed.), *Technologies of the 21st century in jurisprudence: works of the All-Russia scientific-practical conference (Yekaterinburg, May 24, 2019)* (pp. 145–149). Yekaterinburg: Uralskiy gosudarstvenniy yuridicheskiy universitet. (In Russ.).

Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences*, *39*(4), 837–843. https://doi.org/10.1016/j.paid.2005.01.029

Timofeev, A. V. (1978). *Robots and artificial intelligence*. Moscow: Glavnaya redaktsiya fiziko-matematicheskoy literatury izdatelstva "Nauka". (In Russ.).

Vardi, M. (2012). Artificial Intelligence: Past and Future. *Communications of the ACM*, *55*, 5. https://doi.org/10.1145/2063176.2063177

Watkins, R., & Human, S. (2023). Needs-aware artificial intelligence: AI that 'serves [human] needs'. *AI Ethics*, *3*. https://doi.org/10.1007/s43681-022-00181-5

Winfield, A. (2019). Ethical standards in robotics and AI. *Nature Electronics*, *2*, 46–48. https://doi.org/10.1038/s41928-019-0213-6

Author information

Dmitriy V. Bakhteev – Doctor of Law, Associate Professor, Department of Criminology, Ural State Law University named after V. F. Yakovlev, Head of CrimLib.info project group
Address: 21 Komsomolskaya Str., 620137, Ekaterinburg, Russian Federation
E-mail: ae@crimlib.info
ORCID ID: https://orcid.org/0000-0002-0869-601X
ScopusAuthorID:https://www.scopus.com/authid/detail.uri?authorId=57208909117
Web of Science Researcher ID:
https://www.webofscience.com/wos/author/record/ABA-1494-2020
Google Scholar ID: https://scholar.google.ru/citations?user=h0zOOdcAAAAJ
РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=762765

Conflict of interests

The author declare no conflict of interests.

Financial disclosure

The research was not sponsored.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – March 25, 2023
Date of approval – April 24, 2023
Date of acceptance – June 16, 2023
Date of online placement – June 20, 2023