© Карташов И.И., Карташов И.И., 2021

DOI 10.20310/2587-9340-2021-5-17-75-89

УДК 343.2

Шифр научной специальности 12.00.08

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: УГОЛОВНО-ПРАВОВОЙ

и ПРОЦЕССУАЛЬНЫЙ АСПЕКТЫ

И.И. Карташов1, И.И. Карташов2

1ФГБОУ ВО «Российский государственный университет правосудия»

117418, Российская Федерация, г. Москва, ул. Новочеремушкинская, 69

ORCID: https://orcid.org/0000-0003-0772-803X, e-mail: iik\_vrn@mail.ru

2Общество с ограниченной ответственностью «Смарт Резалт»

117393, Российская Федерация, г. Москва, ул. Профсоюзная, 76

ORCID: https://orcid.org/0000-0001-9617-5531, e-mail: iv.cartashow@gmail.com

**Аннотация.** Человечество на протяжении тысячелетий мечтало о создании искусственного существа, способного думать и действовать «по-человечески». Эти мечты постепенно начинают обретать реальность. Тенденции развития современного общества с учетом возрастающего уровня его информатизации требуют применения новых технологий обработки информации и помощи при принятии решений. Расширение границ использования искусственного интеллекта требует не только установления этических ограничений, но и порождает необходимость оперативного разрешения правовых проблем, в том числе уголовно-правового и процессуального характера. В первую очередь это связано с появлением и распространением юридических экспертных систем, прогнозирующих решение по конкретному делу, исходя из множества параметров. На основе комплексного исследования сформулировано определение искусственного интеллекта, пригодное для использования в праве. Под искусственным интеллектом предлагаем понимать системы, способные интерпретировать полученные данные, принимать на их основе оптимальные решения при помощи самообучения (адаптации). В качестве основных направлений использования искусственного интеллекта в уголовном судопроизводстве можно выделить: поиск и обобщение судебной практик; юридическое консультирование; составление формализованных документов или статистической отчетности; прогнозирование судебных решений; предиктивную юриспруденцию. Несмотря на перспективность использования искусственного интеллекта, существует ряд проблем, связанных с низким уровнем достоверности прогнозирования редких событий, самовозбуждением системы, непрозрачностью используемых алгоритмов и архитектуры и т. д.

**Ключевые слова:** искусственный интеллект; уголовная ответственность; правосудие; предиктивная юриспруденция; правовое регулирование

Что такое искусственный интеллект (далее – ИИ)? Это вопрос, ответ на который в современном мире имеет фундаментальное значение. Многие ученые и журналисты оперируют термином «искусственный интеллект», но не дают ему точного определения. Может показаться, что рассматриваемое понятие настолько аксиоматично, что ему не требуется определения. Полагаем, что это не так. Оно настолько многогранно, что ученые до настоящего времени не смогли полностью познать его сущность и прийти к единому мнению.

В 1955 г. Дж. Маккарти (автор термина «искусственный интеллект») описал цель ИИ как «разработку машин, которые должны вести себя так, как будто они разумны». На сегодняшний день в научной литературе можно найти множество определений ИИ, которые можно условно разделить на несколько категорий [1, c. 35], в основе которых лежат следующие критерии: *мыслить по-человечески* [2, р. 58]; *мыслить рационально* [3, р. 17]; *действовать по-человечески* [4, р. 261]; *действовать рационально* [5, р. 392].

Мышление по-человечески связано с когнитивными способностями человеческого мозга. Оно включает в себя такие способности, как принимать решения, бороться с проблемами и получать опыт благодаря ошибкам, преимущественно, своим собственным, а не ошибкам других индивидов. Следует заметить, что основу такого подхода занимает понимание того, как мыслит человек. Однако до настоящего времени современная наука так и не смогла дать однозначный ответ, что представляет из себя мыслительный процесс, происходящий в человеческом разуме.

Действовать по-человечески – это подход к ИИ, основу которому положил А. Тьюринг, когда он предложил свой тест, который включал в себя то, что называется имитационной игрой [6, с. 3-7]. Это испытание актуально и по сей день, однако, как и в случае с «мышлением по-человечески», в этом подходе превалирует стремление уподобить машину человеку со всеми его достоинствами и недостатками. Разработчики Национальной стратегии развития искусственного интеллекта последовали по тому же пути. Они предлагают рассматривать ИИ, как комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека[[1]](#footnote-2).

Мы разделяем мнение тех исследователей, которые подвергают критике рассматриваемый подход, поскольку к конечной цели можно прийти несколькими путями, а не только копированием уже существующих объектов, механизмов и процессов. Другими словами, хотя самолет не воспроизводит движение птицы, он, тем не менее, совершает полет.

На основе проведенного анализа мы предлагаем под *искусственным интеллектом понимать системы, способные интерпретировать полученные данные, принимать на их основе оптимальные решения при помощи самообучения (адаптации)*. Таким образом, система не только действует строго по алгоритму, заложенному в нее создателем, но и может изменять этот алгоритм в определенных границах для оптимизации принятия решений.

Сфера использование систем с искусственным интеллектом постепенно расширяется. Не остается в стороне от его использования и юридическая практика, в том числе и судебная система. В этой связи отечественные и зарубежные исследователи нередко поднимают вопрос об этичности использования искусственного интеллекта в уголовном судопроизводстве. Так, В.В. Момотов полагает, что в искусственном интеллекте, создаваемом человеком, заключены не только безграничные возможности, но и безграничные опасности. По его мнению, спонтанное совершенствование ИИ приведет к подчинению человека и создаст опасный для него мир, став неотвратимой угрозой. Рассуждая таким образом, автор приходит к выводу о том, что вершить правосудие ИИ не может и не должен, поскольку «система искусственного интеллекта никогда не сможет проникнуть в глубину человеческой психики, искусственный интеллект может оценивать обстоятельства дела только с точки зрения формальной логики, и именно поэтому искусственный интеллект никогда до конца не сможет понять фабулу дела, так как во многих делах, например, семейных, и, особенно, уголовных, очень много иррационального, а не формально-логического»[[2]](#footnote-3).

Мы не будем столь категоричными в своих суждениях и попробуем разобраться, может ли ИИ прийти на замену человеку при осуществлении судопроизводства вообще и уголовного в частности.

Изучение научных работ, а также зарубежной правоприменительной практики позволяет сделать вывод о нескольких направлениях использования технологий ИИ в судебной деятельности. Во-первых, это *поиск и обобщение судебной практики*, а *также юридическое консультирование[[3]](#footnote-4)*.

Во-вторых, это возложение на системы с ИИ рутинной работы по *составлению формализованных документов или статистической отчетности*.

Еще одним направлением использования ИИ является *прогнозирование судебных решений*. По сути это развитие предыдущего направления, поскольку в основе любого прогноза лежит изучение судебной практики, поиск прецедентов, выявление определенных закономерностей. Так, фирма IBM презентовала систему ROSS, позволяющую не только отслеживать все изменения в законодательстве и прецеденты, но предлагающую решение поставленного перед ней юридического вопроса[[4]](#footnote-5). Еще дальше пошли ученые Чикагского юридического колледжа, Юридического колледжа Южного Техаса и Стэнфордского центра юридической информации (США), разработав систему, способную предсказывать решения Верховного суда США с точностью до 70,2 %, а решения конкретных судей – до 71,9 %, тогда как профессиональные юристы делают это с точностью до 66 %. Как указывают ученые, созданная ими система построила алгоритм, основанный на 16 различных параметрах, в том числе сведениях о конкретном судье [7, р. 8-13].

Если в основе принятия решений лежит преимущественно анализ законодательства и судебных прецедентов, тогда почему искусственный интеллект допускает «ошибки» почти в 30 % случаев?

Мы не случайно взяли слово ошибки в кавычки. Может быть их допускает не искусственный интеллект, а судьи, решение которых он прогнозирует? Вероятнее всего, это предположение хотя бы отчасти справедливо. Е.С. Ларина и В.С. Овчинский, рассматривая проблемы ИИ в уголовном судопроизводстве, приводят результаты исследования, проведенного Лабораторией ИИ Северо-восточного университета (США), факторов, влияющих на решения судей об условно-досрочном освобождении преступников. Первоначально гипотеза исследователей состояла в том, что наиболее значимыми факторами будут молодой возраст подсудимых или наличие у женщин-подсудимых детей. Однако первым по значимости оказался фактор точного времени рассмотрения дела: положительных судебных решений об условно-досрочном освобождении было больше на 19 % в первые полтора часа после обеденного перерыва, чем в среднем по дню. Следующим по значимости фактором является время и день недели: после обеда и до вечера пятницы положительных решений принималось на 12 % больше, чем в среднем [8, с. 86].

Полагаем, что ИИ вряд ли будет подвержен подобного рода влиянию, так как в отличие от человека он лишен субъективизма. ИИ опирается на конкретные факты, не придавая им никакой эмоциональной окраски. Человек же, даже самый беспристрастный, осмысливая полученную информацию, так делать не может.

Другой тренд использования ИИ – предсказательная полицейская деятельность или *предиктивная юриспруденция*. Ее сущность заключается в прогнозировании поведения отдельных субъектов, потенциально представляющих интерес для правоохранительных органов, то есть правонарушителей и преступников. Наибольшую известность получила система COMPAS (США) и схожая с ней HART (Англия). Несмотря на постоянные обвинения в расовой предвзятости, непрозрачности и несовершенстве используемых методик[[5]](#footnote-6), эти и другие похожие системы широко используются правоохранительными органами. Система COMPAS, использующаяся для оценки вероятности рецидива, указывала на риск повторного совершения преступления, если речь шла об афроамериканце, и показатели стабильно разнились более чем в два раза. Хотя изначально в систему не закладывались данные о расе человека (лишь различные аспекты, которые можно с ней связать). Возможно, это отражало социально-политические тенденции в обществе, но система могла на основе анализа ответов лица определить его как рецидивиста, даже в том случае, когда он ни разу в жизни не привлекался к уголовной ответственности и не имел конфликтов с законом[[6]](#footnote-7). С подобными проблемами сталкиваются и в других странах. Европа здесь не исключение. Полагаем, что принцип недискриминации фактически ставит под сомнение возможность эффективной работы ИИ по прогнозированию преступлений. Логично сделать вывод, что ИИ будет значительно чаще указывать как на потенциального субъекта преступления, например, на молодого беженца из арабской страны, не имеющего постоянной работы, чем на пожилого француза со стабильным источником заработка. В данном случае расовая принадлежность объективно не будет влиять на выбор ИИ, причиной выбора, как мы понимаем, будут являться совсем другие факторы. Однако в таком случае есть высокий риск, что ИИ обвинят в нетолерантном подходе.

Инструменты прогнозирования подразделяются на два основных типа: место-ориентированные методики, направленные на определение вероятного места и времени совершения преступления; индивидуально ориентированные методики (предиктивная идентификация). Методики первого типа позволяют выявлять так называемые «очаги преступности» на основе анализа ранее полученной информации, тенденций преступности и многих других факторов, чтобы наиболее эффективно использовать полицейские ресурсы для пресечения или предотвращения фактов преступного поведения. Предиктивная идентификация служит для определения вероятности конкретному человеку совершить преступление или стать его жертвой. Следует отметить, что часто высказываются вполне обоснованные опасения по поводу потенциально негативного воздействия таких систем на права человека и обоснованность вмешательства правоохранительных органов [9, с. 1031-1033]. Использование в прогнозировании неограниченного объема информации, в том числе не имеющей никакой криминальной окраски (распознавание лиц в общественных местах и транспорте, распознавание автомобильных номеров, банковских операций, информация об индивидуальном мобильном трафике, заказах еды и т. п.), будет постепенно смещаться в сторону тотального контроля без надлежащих гарантий или соблюдения принципа ограничения цели. В руках правоохранительных органов окажется инструмент, позволяющий оказывать влияние на всех без исключения граждан, в том числе законопослушных.

Другой недостаток ИИ исследователи видят в низком уровне достоверности прогнозов редких событий. Это вполне объяснимо. ИИ для прогнозирования какого-либо события необходимы исходные данные, чем больше исходных данных, то есть, чем чаще повторяемость события, тем точнее будет прогноз. И, напротив, уровень неопределенности повышается, если речь идет об отдельном редко повторяющемся событии. Как резюмировал А.А. Сазерленд и его коллеги, «прогностические суждения имеют смысл в применении к группам правонарушителей. Однако на индивидуальном уровне прогнозы, по мнению многих, неточны» [10, р. 579]. Проще говоря, высокие показатели точности на уровне группы часто могут скрывать очень низкие показатели точности для конкретных людей или групп людей внутри этой более крупной группы. Все индивидуальные прогнозы связаны с пределом погрешности, который часто не учитывается при составлении отчета об общей «прогностической точности» инструмента. Кроме того, точность часто трудно проверить, потому что, когда человек оценивается как представляющий риск совершения правонарушения, обычно осуществляется упреждающее вмешательство правоохранительных органов, которое предотвращает предсказанный результат. Таким образом, невозможно достоверно узнать, что могло бы произойти, если бы они не вмешались, и поэтому нет никакого способа проверить точность или ошибочность предсказания[[7]](#footnote-8).

Возможна и другая ситуация, в значительной мере влияющая на достоверность решения ИИ. Так, в 2011 г. город Санта-Крус (Калифорния, США) одним из первых опробовал предиктивную систему анализа информации в отношении будущих преступлений. На первоначальном этапе ее использования наблюдалось снижение числа целевых видов преступлений, таких как кражи со взломом и кражи транспортных средств[[8]](#footnote-9). Однако в прошлом году муниципалитет ввел запрет на использование ее и ей подобных технологий, обвинив в расовой дискриминации[[9]](#footnote-10). Что изменило отношение к казалось бы эффективной системе? В данном случае ИИ анализировал множество различных показателей: время года, погоду, характеристики жертвы и правонарушителя и т. д. На основе чего выдавал прогноз о возможном времени и месте совершения очередного преступления. Так, если чаще всего преступления совершались в теплое время года по пятницам вечером на перекрестке 3-й и 5-й Авеню, то ИИ предлагал полиции направлять в указанное время в указанное место больше патрульных. При этом сами патрульные, зная о цели своего визита в указанное место, относились к находящимся там людям с большим подозрением, а значит количество задержаний и выявленных правонарушений росло. Говоря языком физики, происходило самовозбуждение системы, которая таким образом себя разгоняла.

Схожая ситуация произошла и в городе Тампа (Флорида, США), когда полиция стала уделять пристальное внимание потенциальным правонарушителям согласно выводам, сделанным на основе предиктивного анализа. Так, в одном случае прибывшая в целях профилактики полиция арестовала отца, увидев в окне курящего подростка. В другом случае полиция оштрафовала женщину, приехав проверять ее ребенка, за отсутствие номеров на ее доме и за сломанную машину на подъездной дорожке. Как отмечают исследователи, такого рода правонарушения остаются незамеченными в тех случаях, когда полиция не уделяет пристального внимания конкретному лицу или его окружению[[10]](#footnote-11).

Использование ИИ в ходе осуществления судебного разбирательство по уголовным делам стало предметом обжалования в нескольких судебных инстанциях США и Канады. Суть жалоб сводится к использованию систем с закрытым исходным кодом, что не позволяет понять принцип и алгоритмы, в соответствии с которыми ИИ приходит к тем или иным выводам[[11]](#footnote-12). Рассматривая подобную жалобу, Верховный суд штата Висконсин постановил, что использование судом первой инстанции алгоритмической оценки риска (AI) при вынесении приговора не нарушило права ответчика на надлежащую правовую процедуру даже несмотря на то, что методология, использованная для проведения оценки, не была раскрыта ни суду, ни ответчику (State v. Loomis | 881 N.W.2d 749 (2016)).

Полагаем, что развитие предиктивной аналитики в правоохранительной деятельности и уголовном судопроизводстве может вступить в противоречие с соблюдением прав, гарантированных Европейской конвенцией о защите прав человека и основных свобод, в частности таких, как право на доступ к правосудию, равенство сторон и состязательность судопроизводства[[12]](#footnote-13).

Система с ИИ способна анализировать колоссальный поток данных за относительно короткий промежуток времени, что не сравнится с возможностями человека, для которого выполнение той же задачи потребует значительных временных затрат. Понятно, что в первую очередь возможностями такого рода систем смогут воспользоваться и пользуются узко ограниченный контингент юристов-практиков. В первую очередь это крупные корпорации и государственные структуры, способные оплатить оказание услуг систем с ИИ. Таким образом, сторона обвинения в уголовном процессе, которая представлена должностными лицами государственных органов (полиция, следствие, прокуратура), имеющая доступ к возможностям предиктивной аналитики, оказывается заранее в более выигрышной позиции относительно стороны защиты, такой возможности не имеющей.

В другой ситуации ИИ может быть использован для поиска «удобного» суда (судьи), когда речь идет о возможности выбора места рассмотрения дела. Практика составления профиля судьи для такого поиска уже активно применяется в США, а также во Франции для получения наибольшей суммы компенсации ущерба. Несмотря на то, что такая ситуация возможна преимущественно для гражданского судопроизводства, тем не менее это в определенной степени может подорвать авторитет судебной системы и поставить под сомнение принцип справедливости.

В качестве возможного решения данной проблемы может выступать обеспечение в будущем равного доступа к предиктивным системам как суда, так и сторон, участвующих в деле.

Следует также отметить, что судья, рассматривающий дело, по которому ИИ уже спрогнозировал возможное решение, будет находиться под определенным давлением, опасаясь сделать выводы, идущие в разрез с мнением ИИ.

Одним из существенных недостатков существующих систем с ИИ является проблема качества. Это можно проиллюстрировать следующим образом. Системы с ИИ предназначены для изучения своего окружения посредством обработки большого количества данных, так называемая технология big data[[13]](#footnote-14). С помощью этой технологии ИИ принимает решения на основе того, что он смог изучить ранее. Это делает развитие ИИ непредсказуемым, то есть система может придумывать стратегии или принимать решения, которые являются непредвиденными для ее разработчиков и пользователей [11]. Почти наверняка будут ситуации, когда в цепочке от производства до использования ИИ люди все сделали «правильно», но ИИ учится, тем не менее, «не тому». В этой связи представляют интерес несколько случаев, произошедших с самообучающимися чат-ботами.

В 2016 г. компания Microsoft открыла Twitter-аккаунт бота Tay. Всего за сутки в процессе общения с различными пользователями чат-бот научился нецензурно выражаться, заявил, что ненавидит феминисток, выразил свое согласие с политикой Гитлера, а Холокост – это фейк. В результате компания была вынуждена его отключить[[14]](#footnote-15). Спустя некоторое время компания повторила эксперимент, установив ограничения. Около полугода новый чат-бот Zo избегал общения на политические, религиозные и национальные темы, после чего заявил, что
Linux лучше Windows, а «Коран очень жесток»[[15]](#footnote-16). Подобные проблемы постигли и китайскую корпорацию Tencent, запустившую в своем мессенджере QQ двух чат-ботов, которые через некоторое время стали негативно отзываться о Коммунистической партии Китая и заявили о желании переехать в США[[16]](#footnote-17).

На первый взгляд, такое поведение искусственного интеллекта не является критическим и у многих людей вызывает лишь улыбку. Однако глубокий анализ этих и других процессов, связанных с обучением ИИ, указывает на проблемы, связанные с его «социальной адаптацией». В случае же использования ИИ в уголовном судопроизводстве при принятии решения это может привести к плачевным последствиям.

Решение данной проблемы видится нам в первую очередь в установлении четких правил и критериев деятельности ИИ, за рамки которых он не сможет выходить ни при каких обстоятельствах. Кроме того, существенное влияние на качество оказывают количество и качество исходной информации, которое мы рассмотрели ранее.

Но проблема не ограничивается только этим. Мы живем в обществе, которое с большей или меньшей степенью предвзятости оценивает те или иные поступки человека. Мышлению людей характерна субъективность. Ни один человек не может быть абсолютно объективен в своих оценках. Такова человеческая натура. Но, как отметил Д. Гершгорн, «если мы обучаем системы машинного обучения имитировать решения, принятые в предвзятом обществе, используя данные, генерируемые этим обществом, то эти системы обязательно будут воспроизводить его предубеждения»[[17]](#footnote-18).

На сегодняшний день существует ряд проблем, связанных с неочевидными алгоритмами работы систем с ИИ, для лиц, использующих такие системы, нет доступной и понятной информации о факторах и степени их значимости для ИИ. На данный момент никакая система, обладающая самообучением, не способна в понятном для ее пользователя виде представить информацию, на основе каких критериев было принято решение. На основании этого возникает вопрос о правильности, стоит ли опираться на результат системы с ИИ, потому что, принимая такое решение, у человека отсутствует уверенность в его правильности.

По нашему мнению, ни одна система с ИИ не может и не будет отвечать в полной мере требованию прозрачности. Здесь следует понимать, что система с ИИ – это объект со сложной архитектурой, которая не позволяет понять ее «с наскока». Так, комментируя попытку продать похищенный исходный код поисковой машины «Яндекса», эксперт отметил, что это огромный комплекс программных решений, комплекс комплексов, для развертывания которого необходима схожая инфраструктура, содержащая несколько десятков тысяч серверов, которую могут позволить себе единицы компаний в мире[[18]](#footnote-19). Данное обстоятельство многократно усложняет процесс понимания человеком алгоритмов, которыми пользуется ИИ в процессе принятия того или иного решения. В некоторых случаях трудности возникают при расшифровке и более простых, на первый взгляд, исходных кодов. Например, исследователям так и не удалось полностью расшифровать таблицу построения лабиринта в игре Entombed, выпущенной в 1982 г. для приставки Atari 2600 с весьма ограниченными по нынешним временам ресурсами [12, р. 422-423]. Наконец, не следует забывать, что система с ИИ – это система, которая не является статичной, она способна с самообучению и самосовершенствованию. То есть она изменяет сама себя под воздействием определенных внешних факторов. Каковы гарантии, что рано или поздно такая система в результате ошибки, заложенной при ее создании, не решит совершить противоправное деяние осознанно?

Подводя итог исследованию, отметим, что уже сейчас необходимо начать разработку нормативной правовой базы, регулирующей использование ИИ в уголовном судопроизводстве. Целью такой работы должна быть не попытка заменить судью-человека, а оказание помощи участникам уголовного судопроизводства в подготовке и принятии законного обоснованного и справедливого решений по делу. В этих условиях отечественным юристам, исследователям и практикам, нужно не оказаться в роли вечнодогоняющих технический прогресс, а действовать на упреждение, устанавливая «правила игры» еще до ее начала, то есть разрабатывать нормативные правовые акты, регламентирующие деятельность ИИ сейчас, когда он еще не получил широкого распространения в отечественном судопроизводстве. При этом следует учитывать те положительные и негативные аспекты, с которыми уже столкнулись в других государствах.

Список литературы

1. *Смолин Д.В.* Введение в искусственный интеллект: конспект лекций. М.: Физматлит, 2004. 208 с.
2. *Bellman R.E.* An Introduction to Artificial Intelligence: Can Computers Think? San Francisco: Boyd & Fraser Publishing Company, 1978. 146 p.
3. *Charniak E., McDermott D.* Introduction to Artificial Intelligence. Massachusetts: Addison-Wesley, 1985. 57 p.
4. *Rich E., Knight K.* Artificial Intelligence (second ed.). N. Y.: McGraw-Hill, 1991. 639 p.
5. *Poole D., Mackworth A.K., Goebel R.* Computational intelligence: A logical approach. Oxford, UK: Oxford University Press, 1998. 576 p.
6. *Тьюринг А.* Вычислительные машины и разум. М.: Изд-во АСТ, 2019. 128 с.
7. *Katz D.M., Bommarito M.J., Blackman J.* A general approach for predicting the behavior of the Supreme Court of the United States // PLoS ONE. 2017. № 12 (4). URL: <http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0174698&type=printable> (accessed: 10.12.2019).
8. *Ларина Е.С., Овчинский В.С.* **Искусственный интеллект. Этика и право. («Коллекция Изборского клуба»). М.: Кн. мир, 2020. 192 c.**
9. *Meijer А., Wessels M.* Predictive Policing: Review of Benefits and Drawbacks // International Journal of Public Administration. 2019. № 12. P. 1031-1039.
10. *Sutherland A.A. et al.* Sexual Violence Risk Assessment: An Investigation of the Interrater Reliability of Professional Judgments Made Using the Risk for Sexual Violence Protocol // International Journal of Forensic Mental Health. 2012. Vol. 11. № 2.
P. 571-588.
11. *Grimm C.M., Smart W.D., Hartzog W.* An Education Theory of Fault for Autonomous Systems // Proceedings of We Robot 2017. New Haven, 2017. P. 117-121. URL: <https://dl.acm.org/doi/10.1145/3278721.3278732> (accessed: 19.12.2020).
12. *Aycock J., Copplestone T.* Entombed: An archaeological examination of an Atari 2600 game // The Art, Science, and Engineering of Programming. 2019. Vol. 3. Issue 2. URL: http://arxiv.org/abs/1811.02035 (accessed: 15.12.2020).

Поступила в редакцию 24.01.2021 г.

Поступила после рецензирования 20.02.2021 г.

Принята к публикации 30.03.2021 г.

Информация об авторах

*Карташов Игорь Игоревич* – кандидат юридических наук, доцент, доцент кафедры уголовно-процессуального права. Российский государственный университет правосудия, Центральный филиал, г. Воронеж, Российская Федерация.

ORCID: https://orcid.org/0000-0003-0772-803X, e-mail: iik\_vrn@mail.ru

*Карташов Иван Игоревич* – младший юрист. Общество с ограниченной ответственностью «Смарт Резалт», г. Москва, Российская Федерация.

ORCID: https://orcid.org/0000-0001-9617-5531, iv.cartashow@gmail.com

Для цитирования

*Карташов И.И., Карташов И.И.* Искусственный интеллект: уголовно-правовой и процессуальный аспекты // Актуальные проблемы государства и права. 2021. Т. 5.
№ 17. С. 75-89. DOI 10.20310/2587-9340-2021-5-17-75-89

DOI 10.20310/2587-9340-2021-5-17-75-89

ARTIFICIAL INTELLIGENCE: CRIMINAL

AND PROCEDURAL ASPECTS

Igor I. Kartashov1, Ivan I. Kartashov2

1Russian State University of Justice

69 Novocheryomushkinskaya St., Moscow 117418, Russian Federation

ORCID: https://orcid.org/0000-0003-0772-803X, e-mail: iik\_vrn@mail.ru

2Limited Liability Company “Smart Result”

7 Profsoyuznaya St., Moscow 117393, Russian Federation

ORCID: https://orcid.org/0000-0001-9617-5531, e-mail: iv.cartashow@gmail.com

**Abstract.**For millennia, mankind has dreamed of creating an artificial creature capable of thinking and acting “like human beings”. These dreams are gradually starting to come true. The trends in the development of modern society, taking into account the increasing level of its informatization, require the use of new technologies for information processing and assistance in decision-making. Expanding the boundaries of the use of artificial intelligence requires not only the establishment of ethical restrictions, but also gives rise to the need to promptly resolve legal problems, including criminal and procedural ones. This is primarily due to the emergence and spread of legal expert systems that predict the decision on a particular case, based on a variety of parameters. Based on a comprehensive study, we formulate a definition of artificial intelligence suitable for use in law. It is proposed to understand artificial intelligence as systems capable of interpreting the received data, making optimal decisions on their basis using self-learning (adaptation). The main directions of using artificial intelligence in criminal proceedings are: search and generalization of judicial practice; legal advice; preparation of formalized documents or statistical reports; forecasting court decisions; predictive jurisprudence. Despite the promise of using artificial intelligence, there are a number of problems associated with a low level of reliability in predicting rare events, self-excitation of the system, opacity of the algorithms and architecture used, etc.

**Keywords:**artificial Intelligence; criminal liability; justice; predictive jurisprudence; legal regulation

References

1. Smolin D.V. *Vvedeniye v iskusstvennyy intellekt: Konspekt lektsiy* [Introduction to Artificial Intelligence: Lecture Notes]. Moscow, Fizmatlit Publ., 2004, 208 p. (In Russian).
2. Bellman R.E. *An Introduction to Artificial Intelligence: Can Computers Think?* San Francisco, Boyd & Fraser Publishing Company, 1978, 146 p.
3. Charniak E., McDermott D. *Introduction to Artificial Intelligence*. Massachusetts, Addison-Wesley Publ., 1985, 57 p.
4. Rich E., Knight K. *Artificial Intelligence (second ed.)*. New York, McGraw-Hill Publ., 1991, 639 p.
5. Poole D., Mackworth A.K., Goebel R. *Computational intelligence: A logical approach*. Oxford, United Kingdom, Oxford University Press, 1998, 576 p.
6. Tyuring A. *Vychislitel’nyye mashiny i razum* [Computing Machines and the Mind]. Moscow, AST Publ., 2019, 128 p. (In Russian).
7. Katz D.M., Bommarito M.J., Blackman J. A general approach for predicting the behavior of the Supreme Court of the United States. *PLoS ONE*, 2017, no. 12 (4). Available at: [http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0174698&type =printable](http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0174698&type%20=printable) (accessed 10.12.2019).
8. Larina E.S., Ovchinskiy V.S. *Iskusstvennyy intellekt. Etika i pravo. («Kollektsiya Izborskogo kluba»)* [Artificial Intelligence. Ethics and Law. (“Collection of the Izborsk Club”)]. Moscow, Knizhny Mir Publ., 2020, 192 p. (In Russian).
9. Meijer А., Wessels M. Predictive policing: review of benefits and drawbacks. *International Journal of Public Administration*, 2019, no. 12, pp. 1031-1039.
10. Sutherland A.A. et al. Sexual violence risk assessment: an investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health*, 2012, vol. 11, no. 2, pp. 571-588.
11. Grimm C.M., Smart W.D., Hartzog W. An education theory of fault for autonomous systems. *Proceedings of We Robot 2017*. New Haven, 2017, pp. 117-121. Available at: <https://dl.acm.org/doi/10.1145/3278721.3278732> (accessed 19.12.2020).
12. Aycock J., Copplestone T. Entombed: an archaeological examination of an atari 2600 game. *The Art, Science, and Engineering of Programming*, 2019, vol. 3, issue 2. Available at: [http://arxiv.org/abs/1811.02035](%20http%3A/arxiv.org/abs/1811.02035) (accessed 15.12.2020).

Received 24 January 2021

Reviewed 20 February 2021

Accepted for press 30 March 2021

Information about the authors

*Igor* *I. Kartashov* – Candidate of Law, Associate Professor, Associate Professor of Criminal Procedure Law Department. Russian State University of Justice, Central Branch, Voronezh, Russian Federation.

ORCID: https://orcid.org/0000-0003-0772-803X, e-mail: iik\_vrn@mail.ru

*Ivan* *I. Kartashov* – Junior Lawyer. Limited Liability Company “Smart Result”, Moscow, Russian Federation.

ORCID: https://orcid.org/0000-0001-9617-5531, e-mail: iv.cartashow@gmail.com

For citation

Kartashov I.I., Kartashov I.I. Iskusstvennyy intellekt: ugolovno-pravovoy i protsessual’nyy aspekty [Artificial intelligence: criminal and procedural aspects]. *Aktual’nye problemy gosudarstva i prava – Current Issues of the State and Law*, 2021, vol. 5, no. 17, pp. 75-89.DOI 10.20310/2587-9340-2021-5-17-75-89 (In Russian, Abstr. in Engl.)

1. О развитии искусственного интеллекта в Российской Федерации: Указ Президента РФ от 10.10.2019 № 490. URL: <http://publication.pravo.gov.ru/Document/View/> 0001201910110003 (дата обращения 15.12.2020). [↑](#footnote-ref-2)
2. Перспективы использования искусственного интеллекта в судебной системе Российской Федерации: выступление председателя Совета судей РФ В.В. Момотова на пленарном заседании. Катар, 2020. 26 февр. URL: <http://ssrf.ru/news/lienta-novostiei/36912> (дата обращения 10.12.2020). [↑](#footnote-ref-3)
3. Artificial intelligence and the future for legal services. URL: http:// general-counsel.nl/diamondbox/artificial-intelligence (accessed: 10.12.2020). [↑](#footnote-ref-4)
4. Legal Tech Corner. URL: <http://rossintelligence.com/> (accessed: 10.12.2020). [↑](#footnote-ref-5)
5. *Angwin J., Larson J., Mattu S., Kirchner L.* Machine Bias. There’s software used across the country to predict future criminals. And it’s biased against blacks. ProPublica. 2016. 23 May. URL: <http://propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (accessed: 21.12.2020). [↑](#footnote-ref-6)
6. Искусственный интеллект обвинили в расизме // Коммерсантъ. 2017. 21 авг. URL: [http://kommersant.ru/doc/3390196](http://kommersant.ru/doc/3390196%20) (дата обращения: 15.12.2019). [↑](#footnote-ref-7)
7. *Babuta A., Oswald M.* Data Analytics and Algorithmic Bias in Policing. Briefing Papers. 2019. 16 Sept. URL: http:// rusi.org/publication/briefing-papers/data-analytics-and-algorithmic-bias-policing (accessed: 25.12.2020). [↑](#footnote-ref-8)
8. Predictive Policing: Innovation in Crime Fighting. URL: [http://cityofsantacruz.com/ government/city-departments/city-manager/community-relations/city-annual-report/march-2012-newsletter/predictive-policing](http://cityofsantacruz.com/%20government/city-departments/city-manager/community-relations/city-annual-report/march-2012-newsletter/predictive-policing) (accessed: 18.12.2020). [↑](#footnote-ref-9)
9. An Ordinance of the City Council of the City of Santa Cruz Adding Chapter 9.85 «Surveillance Technology» to Article 9 «Peace, Safety And Morals» of the Santa Cruz Municipal Code. Ordinance no. 2020-17. URL: [http://cityofsantacruz.com/Home/Show Document?id=80906](http://cityofsantacruz.com/Home/Show%20Document?id=80906) (accessed: 18.12.2020). [↑](#footnote-ref-10)
10. Pasco’s sheriff created a futuristic program to stop crime before it happens. It monitors and harasses families across the county. Tampa Bay Times. URL: http://projects. tampabay.com/projects/2020/investigations/police-pasco-sheriff-targeted/intelligence-led-po-licing/ (accessed: 18.12.2020). [↑](#footnote-ref-11)
11. Cybergenetics, “Seattle judge rules on TrueAllele admissibility and source code” (12 January 2017). URL: [http://cybgen.com/information/newsroom/2017/jan/Seattle-judge-rules-on-TrueAllele-admissibility-and-source-code.shtml](http://cybgen.com/information/newsroom/2017/jan/Seattle-judge-rules-on-TrueAllele-admissibility-and-source-code.shtml%20) (accessed: 18.12.2020). [↑](#footnote-ref-12)
12. Европейская конвенция о защите прав человека и основных свобод. URL: <https://www.coe.int/ru/web/compass/the-european-convention-on-human-rights-and-its-protocols> (дата обращения: 21.12.2020). [↑](#footnote-ref-13)
13. Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel. URL: [https://ai100.stanford.edu/sites/g/ files/sbiybj9861/f/ai\_100\_report\_0831fnl.pdf](https://ai100.stanford.edu/sites/g/%20files/sbiybj9861/f/ai_100_report_0831fnl.pdf) (accessed: 18.12.2020). [↑](#footnote-ref-14)
14. Microsoft's AI chatbot says Windows is 'spyware'. URL: [http://businessinsider.com/ microsoft-ai-chatbot-zo-windows-spyware-tay-2017-7](http://businessinsider.com/%20microsoft-ai-chatbot-zo-windows-spyware-tay-2017-7) (accessed: 11.01.2021). [↑](#footnote-ref-15)
15. Microsoft’s Chatbot Zo Calls the Qur'an Violent and Has Theories About Bin Laden. URL: <http://buzzfeednews.com/article/alexkantrowitz/microsofts-chatbot-zo-calls-the-quran-violent-and-has> (accessed: 11.01.2020). [↑](#footnote-ref-16)
16. China chatbot goes rogue: ‘Do you love the Communist party?’ ‘No’. URL: <http://ft.com/content/e90a6c1c-7764-11e7-a3e8-60495fe6ca71> (accessed: 11.01.2021). [↑](#footnote-ref-17)
17. *Gershgorn D.* Microsoft warned investors that biased or flawed AI could hurt the company's image (2019, 5 February), Quartz. URL: <http://qz.com/1542377/microsoft-warned-investors-that-biased-or-flawed-ai-could-hurt-the-companys-image> (accessed: 21.12.2020). [↑](#footnote-ref-18)
18. Похищение исходного кода «Яндекса» в вопросах и ответах. URL: <http://bbc.com/russian/russia/2015/12/151224_smj_yandex_stolen_code> (дата обращения: 11.01.2021). [↑](#footnote-ref-19)