Алгоритмы и искусственный интеллект сквозь призму прав человека

ТАЛАПИНА Эльвира Владимировна, главный научный сотрудник Института государства и права Российской академии наук, доктор юридических наук, доктор права (Франция)

Россия, 119019, г. Москва, ул. Знаменка, 10

E-mail: comparative_law@igpran.ru

В программных актах по развитию искусственного интеллекта затрагиваются перспективы его использования в государственно-управленческой деятельности, что является новой научной и практической проблемой. Лежащие в основе функционирования искусственного интеллекта алгоритмы известны праву, причем праву частному, только как объекты правовой защиты. В публичном праве актуальна задача теоретического осмысления технологического обновления с учетом устоявшихся правовых ценностей, одна из которых — фундаментальные права человека. При помощи традиционных методов юридического исследования автором проведен анализ российской и зарубежной юридической доктрины, судебной практики, примеров из правоприменительной практики в области подходов к регулированию и использованию технологий в публично-правовых отношениях, что позволило выявить проблемы, с которыми может столкнуться система защиты прав человека.

Идея алгоритма как совокупности инструкций, позволяющих решить проблему и достичь определенного заданного результата, отвечает концепции надлежащего, качественного государственного управления, поскольку результат управленческой деятельности можно заложить в алгоритм принятия управленческого решения. Наибольшая эффективность алгоритмов прогнозируется в тех сферах, где нет дискреционных полномочий, что характерно для сферы оказания государственных услуг. Однако автоматизация управленческих решений несет и возможные риски.

В публичном праве сфера применения алгоритмов (точнее, автоматизированной обработки данных) ограничена персональными данными, угроза которым является одним из рисков при автоматизированном принятии решений. Технология обработки данных состоит в применении к конкретному лицу неких заданных критериев. Прямая угроза фундаментальным правам человека содержится в потенциальной дискриминационности этих критериев. Право на недискриминацию выводится из принципа равенства. Однако формальное юридическое равенство означает также возможность применения общей нормы к конкретному случаю с учетом индивидуальных обстоятельств. Алгоритм же беспристрастно обеспечивает равенство арифметическое. Более того, искусственный интеллект, обучаемый на основании реальных и неидеальных человеческих решений, может воспринять дискриминацию в качестве «нормального» элемента алгоритма. Выявленные проблемы нуждаются в детальной проработке и последующем нормативном решении.

Ключевые слова: алгоритм, искусственный интеллект, дискриминация, персональные данные, равенство, государственные услуги.

Для цитирования: Талапина Э. В. Алгоритмы и искусственный интеллект сквозь призму прав человека // Журнал российского права. 2020. № 10. С. 25—39. DOI: 10.12737/jrl.2020.118

Амбициозные планы человечества по использованию цифровых технологий во всех сферах жизнедеятельности впечатляют мас-

штабами и футуристическими зарисовками, но не стоит забывать, что все это будет происходить непосредственно в нашем обществе.

В обществе со сложившейся правовой культурой, традициями и ожиданиями. Было бы несомненной ошибкой концентрироваться лишь на радужных перспективах, игнорируя риски, которые таят цифровые технологии. Перед публичным правом стоит задача осмысления технологического обновления, с тем чтобы уберечь при этом устоявшиеся правовые ценности, одну из которых представляют собой фундаментальные права человека. Анализ российской и зарубежной юридической доктрины, судебной практики, примеры из правоприменительной практики позволяют выявить проблемы, с которыми может столкнуться система защиты прав человека, и подумать о путях их преодоления.

Право и алгоритм. Прежде всего отметим, что алгоритмы существовали задолго до того, как получили свое название. Они порождены необходимостью находить быстрые и эффективные решения посредством трансформации данных при помощи точных указаний, применяющихся поэтапно. В принципе обычный кулинарный рецепт — это уже алгоритм, который содержит пошаговую инструкцию для достижения определенного результата. Правовая норма — тоже алгоритм, содержащий набор данных, условий их применения в заданной последовательности для достижения некоего результата (решения определенной проблемы). Таким образом, право и алгоритм крайне близки по своей сущности. Если отталкиваться от классических определений — право как система норм, регулирующих отношения между людьми в обществе, и алгоритм как совокупность инструкций, разрешающих проблему, — то понятие алгоритма оказывается более широким. В этом смысле право можно рассматривать как специфический алгоритм (юридический), объединяющий совокупность инструкций, разрешающих проблемы, воз-

никающие из отношений людей в обществе¹.

Но развитие информационных технологий, успехи науки информатики прочно связали понятие алгоритма с информационным обеспечением. Поэтому, используя слово «алгоритм», мы переносимся в область программирования, а право относится к алгоритму как к объекту регулирования.

Правовой охране подлежат не сами алгоритмы, а программы. Программы для ЭВМ в целях охраны авторским правом, в силу указания ст. 1261 ГК РФ и в соответствии с закрепленными на международном уровне подходами условно приравниваются к произведениям литературы. Аналогичное правило предусмотрено в п. 1 ст. 10 Соглашения по торговым аспектам прав интеллектуальной собственности, согласно которому компьютерные программы, которые могут быть выражены на любом языке и в любой форме, включая исходный текст и объектный код, подлежат авторско-правовой охране как литературные произведения. Однако необходимо учитывать специфику программ для ЭВМ, которая вытекает из их предназначения — обеспечивать функционирование компьютерных устройств, осуществление определенных алгоритмов и процессов, достижение результатов, имеющих, по сути, технический характер. Перед литературными произведениями такая задача не ставится².

Алгоритм как последовательность операций является основным компонентом программы, отражая ту главную идею, согласно которой должен работать компьютер. Однако в соответствии с п. 5 ст. 1259 ГК РФ авторские права не распространяются на

¹ См.: Duclercq J.-B. Le droit public à l'ère des algorithmes // *Revue du droit public*. 2017. No. 5. P. 1431.

² См.: Близнец И. А., Леонтьев К. Б., Кубышкин А. В. Правовая охрана авторских и смежных прав в Российской Федерации. М., 2017. С. 146—147.

идеи, концепции, принципы, методы, процессы, системы, способы, решения технических, организационных или иных задач, открытия, факты, языки программирования. В итоге алгоритм лишен правовой защиты; охраняется лишь его реализация в виде последовательности операций и действий над этими операциями³.

Данный подход имеет европейские корни. Согласно Директиве 91/250/СЕ от 14 мая 1991 г. о юридической защите компьютерных программ такая защита предоставлялась только программам, но не идеям и принципам, лежащим в основе алгоритма. Это правило сохранилось и в новой Директиве 2009/24 от 23 апреля 2009 г. Суд Европейского Союза подтвердил изложенную позицию, указав, что защите авторским правом подлежат код-источник (исходный текст) и объектный код компьютерной программы⁴.

Между тем не только в зарубежных странах, но и в России на практике осуществляется патентование алгоритмов (в качестве изобретений (способа))⁵. Программные алгоритмы могут быть запатентованы в качестве изобретений в виде технических решений, относящихся к способу (п. 1 ст. 1350 ГК РФ). Но для этого разработчикам необходимо добиться соответствия решения требованию изобретательского уровня и определить ожидаемый материально-технический результат, который предполагается достичь за счет данного решения⁶.

³ См.: Российское гражданское право: в 2 т. Т. 1: Общая часть. Вещное право. Наследственное право. Интеллектуальные права. Личные неимущественные права: учебник / отв. ред. Е. А. Суханов. М., 2011.

⁴ СЈЕU, 22.12.2010, С-39/09; СЈЕU, 2.05.2012, С-406-10.

⁵ См.: Цивилистическая концепция интеллектуальной собственности в системе российского права: монография / под общ. ред. М. А. Рожковой. М., 2018.

⁶ См.: Ворожеевич А. С., Заиченко Е. В., Кирсанова Е. Е. и др. Современные информа-

Таким образом, алгоритмы хорошо знакомы праву в качестве объектов регулирования, но праву гражданскому, частному. Каковы же перспективы алгоритма в праве публичном?

Алгоритмы в публичном праве. Широкое распространение информационно-коммуникационных технологий, особенно цифровых технологий, их применение в государственном управлении дали основания для введения алгоритмов в поле публичного права. Любопытно, что понятие «алгоритм» в нормативных правовых актах почти не употребляется в его техническом значении, а используется для обозначения порядка действий (например, Алгоритм (порядок) взаимодействия заинтересованных органов при выявлении противоправного контента в сети Интернет, принятый Роскомнадзором).

В Европе, кажется, тоже избегают понятия «алгоритм». С 1981 г. используется термин «автоматизированная обработка данных», который напоминает алгоритм по содержанию (совокупность действий): автоматизированная обработка включает следующие операции, осуществляемые полностью или частично с помощью автоматизированных средств: хранение данных, осуществление логических и (или) арифметических операций с этими данными, их изменение, уничтожение, поиск или распространение (Конвенция Совета Европы о защите физических лиц при автоматизированной обработке персональных данных 1981 г.).

Нетрудно заметить, что сфера применения автоматизированной обработки данных ограничена персональными данными. То есть публично-правовая защита сконцентрирована на человеке с его частной жизнью и персональными данными. Собственно, так и сформировалось право на защиту персональных дан-

ционные технологии и право: монография / отв. ред. Е. Б. Лаутс. М., 2019. С. 208—233.

ных и была заложена традиция «сопротивления» индивидуальным автоматизированным решениям. Так, Регламент № 2018/1725 Европейского парламента и Совета Европейского Союза «О защите физических лиц при обработке персональных данных» (п. 43) определяет, что субъект данных должен иметь право не подчиняться решению, которое может включать принятие мер, оценивающих личные аспекты, относящиеся к нему или к ней, которое основывается исключительно на автоматизированной обработке и результатом которого являются юридические последствия в отношении него или нее или которое сходным образом может значительно повлиять на него или на нее (например, практика электронного подбора кадров без вмешательства человека).

В русле европейских правовых позиций выстроен и российский закон. Согласно ст. 16 Федерального закона от 27 июля 2006 г. № 152-ФЗ «О персональных данных» субъект персональных данных имеет право отказаться от принятия в отношении него решения на основании исключительно автоматизированной обработки персональных данных. К признакам такого решения относится то, что оно порождает юридические последствия в отношении субъекта персональных данных или иным образом затрагивает его права и законные интересы.

При этом имеется в виду самый широкий круг решений — как управленческих, так и любых иных. Обработка персональных данных, осуществляемая в частном секторе, порождает вопросы при заключении и исполнении «договора в онлайн-среде в полностью автоматизированном режиме, например при размещении заказа на приобретение цифрового контента, который становится доступным для загрузки по факту оплаты»⁷.

⁷ Савельев А. И. Научно-практический постатейный комментарий к Федеральному закону «О персональных данных». М., 2017.

Но вернемся к публичному праву, регламентирующему управленческие решения, принимаемые органами власти в отношении граждан на основе автоматизированной обработки данных. Идея алгоритма как совокупности инструкций, позволяющих решить проблему и достичь определенного заданного результата, как нельзя лучше отвечает концепции надлежащего, качественного государственного управления, нацеленного на достижение заранее заданных результатов⁸, — ведь результат управленческой деятельности можно легко заложить в алгоритм принятия управленческого решения. Это дает основания полагать, что в ближайшие годы государство активно займется подобной автоматизацией управленческих решений, а потому необходимо заранее изучить возможные риски.

Прежде всего следует понять, в каких сферах автоматизация может произойти в первую очередь. Поскольку алгоритм — последовательность заранее определенных действий, их наибольшая эффективность прогнозируется в сферах, где не приветствуется широта действий органов власти, иными словами, где нет (и не должно быть) дискреционных полномочий. Эта логика нам хорошо знакома по мероприятиям административной реформы, заложившей основы для предсказуемости действий органов власти. С одной стороны, органы власти «связывались» административными регламентами как шаблонами (алгоритмами) действий, а с другой стороны, свобода дискреционного усмотрения сокращалась на уровне регламентирующих актов путем анти-

⁸ Подробнее о качественном государственном управлении, ориентированном на достижение конкретных результатов, см.: Южаков В. Н., Талапина Э. В., Добролюбова Е. И., Тихомиров Ю. А. Инициативный проект закона об обеспечении качества государственного управления. М., 2020.

коррупционной экспертизы. В итоге гражданин или юридическое лицо при наличии необходимых оснований, подтвержденных документами, получили возможность рассчитывать на предсказуемый, ожидаемый результат (а не сюрприз) своих административных демаршей. Интерфейс такого взаимодействия власти с населением хорошо знаком каждому из нас по portalу государственных услуг, где по ячейкам разложены разделы: требуемые документы, куда обратиться, сроки и результаты. Таким образом, первый кандидат для автоматизации принятия решений — сфера оказания государственных услуг.

Казалось бы, автоматизация принятия решений в области государственных услуг несет пользу для всех — для государства и граждан: удобство, быстрота, экономия. Однако за этим фасадом скрывается та цена, которую граждане платят за удобство, — их персональные данные. Известно, как легко эти данные «утекают» в Сеть и используются в самых разных, в том числе криминальных, целях⁹. Соответственно, угроза персональным данным является одним из рисков при автоматизированном принятии решений. Но не единственным. Алгоритмические процессоры (в частности, технологии микротаргетинга) используются для манипулирования социальным и политическим поведением, что вызывает озабоченность Комитета министров Совета Европы¹⁰. Кстати, применение

алгоритмов именно в деятельности публичной власти расценивается Рекомендацией СМ/Рес (2020)1 Комитета министров Совета Европы о влиянии алгоритмических систем на права человека как наиболее угрожающее правам человека, поскольку у людей часто нет возможности уклониться от вступления в такого рода отношения с государством.

Выяснить еще одну публично-правовую угрозу поможет понимание того, каким образом производится обработка данных.

Анализ многочисленных персональных данных с целью создания индивидуального или группового профиля получил название профайлинга. Такая селекция и отнесение к группе может основываться на различных критериях или социально-экономических характеристиках. Например, выявить наркомана поможет профиль, складываемый из следующих данных: возраст до 40 лет, большое количество нештатных ситуаций на работе, частые отсутствия на работе без уважительной причины. Сходные техники используются таможенниками при выявлении наркотрафика: под подозрение попадают пассажиры с малым количеством багажа, следующие в определенном направлении, совершающие много транзитных рейсов, оплатившие билет наличными.

К обработке данных относят в том числе скоринг — оценку разных профилей по определенному критерию, майнинг — процесс поиска корреляций. Эти понятия зародились еще до бурного развития Интернета в так называемых офлайн-отношениях. К собственно интернетовской эпохе относится stalking (слежка) — отслеживание в Интернете (когда необходимые данные добываются путем поиска следов в Интернете).

Все эти действия по обработке данных основаны на применении к конкретному лицу неких заданных критериев. Как любая технология,

⁹ Отдельную проблему представляет собой хранение государством этих данных, в том числе в «частных хранилищах», расположенных в юрисдикции одного конкретного государства (подробнее см.: Хабриева Т. Я., Черногор Н. Н. Право в условиях цифровой реальности // Журнал российского права. 2018. № 1. С. 95).

¹⁰ См.: Ковлер А. И. Права человека в цифровую эпоху // Бюллетень Европейского суда по правам человека. Российское издание. 2019. № 6. С. 146—150.

эти действия универсальны, т. е. могут использоваться как в частноправовых, так и публично-правовых целях. Когда такие операции становятся частями алгоритма, то применяются уже без вмешательства человека. Но описанная технология обработки данных содержит прямую угрозу фундаментальным правам человека, поскольку заданные в алгоритме критерии могут быть дискриминационными.

Как быть с правами человека?

Принцип всеобщего равенства является главным принципом правового государства. Если нет равенства, права и свободы теряют смысл — ведь тогда ими можно наделять одних и лишать их других. В равенстве выделяются несколько компонентов — равенство перед законом, судом, государственными органами, равенство в смысле притязания на государственную помощь. Впрочем, это не означает, что законодатель не может по-разному регулировать различные ситуации, и даже в некоторых случаях поступиться равенством во имя публичных интересов (как, например, произошло во время пандемии коронавируса).

Разумеется, формальное равенство нарушается в повседневной жизни, но для его обеспечения выстроена вся система правового государства. В эпоху алгоритмов озвучивается надежда на то, что они облегчат бюрократическую работу и за счет своей технической нейтральности обеспечат равенство в публичной деятельности.

Но так ли все однозначно? Смысл юридического равенства не только в том, что все подчинены закону одинаково, но и в том, что общая норма применяется к частному случаю, индивидуализированно, т. е. с учетом всевозможных обстоятельств. Особенно наглядно это демонстрирует институт юридической ответственности, для которого индивидуализация наказания (ответственности) обретает исключительное значение в смысле достижения справедливо-

сти. В индивидуализации — смысл уголовного наказания, административных санкций, гражданско-правовой ответственности. Управленческую практику индивидуализирует мотивация административных решений. Право на надлежащее государственное управление включает обязанность государственных органов мотивировать свои решения (ст. 41 Хартии Европейского Союза об основных правах 2007 г.). А алгоритм формализует мотивацию решений, делает ее стереотипной.

Существует большая опасность, что алгоритм не охватит все эти нюансы. Если алгоритм обеспечивает равенство, то речь идет скорее о равенстве арифметическом, о беспристрастном рассмотрении частных ситуаций, но без учета особенностей каждого случая. Получается, что алгоритм не обеспечивает юридическое равенство¹¹. Верно замечено, что правоприменение по своей природе далеко не всегда способно осуществляться в рамках строгих алгоритмов. В праве существуют пробелы, а любое социальное взаимодействие бывает многоаспектным, допускает различные и порой неоднозначные оценки¹².

За «парадной» стороной равенства как принципа стоит необходимость равного обращения со всеми, или запрет дискриминации.

Многочисленные правовые акты, прежде всего международные (Всеобщая декларация прав человека, Конвенция Совета Европы о защите прав человека и основных свобод, Хартия ЕС об основных правах), подчеркивают связку «равенство —

¹¹ См.: Boudinar-Zabaleta A. Algorithmes et lignes directrices. Réflexions sur la codification automatisée des motifs des décisions administratives // Droit administratif. 2019. No. 4. P. 18.

¹² См.: Пашенцев Д. А. Особенности правоприменения в условиях цифровизации общественных отношений // Вестник Санкт-Петербургского университета. Право. 2020. № 1. С. 42.

недискриминация». При этом тексты определения дискриминации не содержат, говорится только об обеспечении пользования правами и свободами без какой бы то ни было дискриминации по признаку пола, расы, цвета кожи, языка, религии, политических или иных убеждений, национального или социального происхождения, принадлежности к национальным меньшинствам, имущественного положения, рождения или по любым иным признакам.

Отметим, что ст. 14 Европейской конвенции о защите прав человека и основных свобод запрещает дискриминацию только в отношении тех прав и свобод, которые гарантированы в Конвенции и в Дополнительных протоколах к ней. В этом смысле ст. 14 Конвенции не является самостоятельной гарантией¹³. Чтобы расширить действие ст. 14 и установить общий запрет дискриминации, 4 ноября 2000 г. был принят Протокол № 12 к Конвенции (который Россией подписан, но не ратифицирован). Однако этот общий запрет дискриминации распространен только на власть — «никто не может быть подвергнут дискриминации со стороны каких бы то ни было публичных властей».

Понимание дискриминации и права на недискриминацию выводится из судебной практики. Так, в решениях Суда Европейского Союза сформировался подход к дискриминации как к применению разных правил в сопоставимых ситуациях или применению одного правила к разным ситуациям¹⁴. ЕСПЧ указал, что понятие дискриминации обычно включает случаи, когда

лицо или группа лиц испытывают без надлежащего обоснования менее благоприятное обращение, чем другие лица или группы лиц¹⁵. Это прямая дискриминация. Но есть еще и косвенная дискриминация, которая впервые была признана в деле «Тлимменос против Греции»¹⁶, где Суд установил нарушение ст. 14 Конвенции в сочетании со ст. 9 в связи с тем, что заявителю было отказано в возможности работать бухгалтером, поскольку ранее он был осужден за отказ от службы в армии по религиозным мотивам. По мнению Суда, дискриминация состояла в том, что греческое законодательство не проводило различие между теми, кто осужден по причине религиозных убеждений, и теми, кто осужден по иным причинам (§ 48). Тем самым имела место косвенная дискриминация, поскольку единые требования применялись ко всем лицам одинаково, тогда как ситуация заявителя значительно отличалась от других лиц. В целом косвенную дискриминацию можно также охарактеризовать как отказ в особом обращении¹⁷ при наличии на то оснований.

Случаи дискриминации, рассмотренные судами, редко касались автоматизированной обработки данных. Но легко спрогнозировать, что алгоритмы и особенно искусственный интеллект способны так же дискриминировать людей, как и люди.

Алгоритм, помноженный на искусственный интеллект. Алгоритмы лежат в основе многих технологий — искусственный интеллект, распределенный реестр, поддержка принятия решений, Интернет, роботы. Напомним, что привлечение искусственного интеллекта к при-

¹³ См.: Бартенев Д. Г. Запрет дискриминации: практический обзор подходов Европейского суда по правам человека // Международное правосудие. 2019. № 1. С. 43—66.

¹⁴ CJEU, Racke v. Hauptzollamt Mainz, C-283/83; CJEU, Finanzamt Köln-Altstadt v. Schumacker, C-279/93; CJEU, Kingdom of Spain v. Council of the European Communities, C-203/86.

¹⁵ Eur. Court H. R. Abdulaziz, Cabales and Balkandali v. the United Kingdom. Applications nos. 9214/80, 9473/81, 9474/81. Judgment of 28 July 1995. § 82.

¹⁶ Eur. Court H. R. Thlimmenos v. Greece. Application no. 34369/97. Judgment of 6 April 2000.

¹⁷ См.: Бартенев Д. Г. Указ. соч. С. 43—66.

нятию решений не связано с цифровой эпохой — первые экспертные системы (юридические и медицинские) на основе искусственного интеллекта появились в 1970-х гг.¹⁸

Почему возникла идея не просто автоматически обрабатывать данные, а использовать искусственный интеллект? Система оценок позволяет квалифицировать склонность, но не помогает понять точно специфику, которая лежит в основании этой склонности. А смысл правильного решения — в понимании специфики и контекста феномена¹⁹. Глобальность статистического подхода и скоринга предопределила использование технологического подхода в виде искусственного интеллекта.

В литературе называются следующие сущностные признаки искусственного интеллекта: способность автономной работы и обмена данными с окружающей средой; способность обучаться на основе полученного опыта и в процессе взаимодействия с окружающей средой; способность адаптировать свои решения под изменяющиеся обстоятельства²⁰. Поскольку эти компоненты весьма напоминают свойства человеческого мозга, нет никаких объективных причин для того, чтобы искусственный интеллект избежал риска вовлечения в дискриминацию. Более того, обучаемый на основе многочисленных данных в сети, происходящих от обычных людей, легко становится и расистом, и фашистом, чему есть масса приме-

ров. Принятие решений — это когнитивный процесс, который приводит к окончательному выбору из ряда альтернативных вариантов. Но естественная ограниченность людей в исследовании приводит к субъективным решениям, оказывающимся дискриминационными по отношению к «невыбранным». Искусственный интеллект, обучаемый на основании реальной коллекции человеческих решений, в большинстве своем неидеальных, может воспринять дискриминацию в качестве «нормального» элемента алгоритма. Причем дискриминация посредством алгоритма — систематическая, что на деле намного вреднее, нежели разовый людской субъективизм²¹.

Поскольку все зависит от того, каким алгоритмом руководствуется искусственный интеллект, в науке формулируется задача достижения беспристрастности классификации, положенной в основу алгоритма. Самый простой путь — усовершенствовать классификатор, убрав самые «чувствительные» параметры, наиболее очевидно свидетельствующие о дискриминации. Обратимся к примерам из «автоматизированной» кадровой политики частных компаний, которая применяется и к конкурсному набору на государственную службу. Найм на работу основывается на отборе «лучших» параметров кандидатов, что в дальнейшем способно привести к дискриминации (если в качестве ориентира избраны лучшие работники с наиболее высокой заработной платой за определенный период, которые все оказались мужчинами, женщины-претенденты будут отсекаются автоматически). Такой, казалось бы, нейтральный параметр, как

¹⁸ См.: *Bourcier D.* De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? // *Droit et société*. 2001. No. 3. P. 853.

¹⁹ См.: *Huret A., Huet J.-M.* L'intelligence artificielle au service du marketing // *L'Expansion Management Review*. 2012. No. 3. P. 20.

²⁰ См.: *Малышкин А. В.* Интегрирование искусственного интеллекта в общественную жизнь: некоторые этические и правовые проблемы // *Вестник Санкт-Петербургского университета. Право*. 2019. № 3. С. 446.

²¹ См.: *Kamiran F., Žliobaitė I., Calders T.* Quantifying explainable discrimination and removing illegal discrimination in automated decision making // *Knowledge and Information Systems*. 2013. No. 35. P. 613—644. DOI: 10.1007/s10115-012-0584-8.

место жительства, может косвенно указывать на национальность или расу (не секрет, что во многих городах мира существуют национальные кварталы). Адрес проживания может также свидетельствовать о социальном положении лица, его достатке. Сложность состоит еще и в том, что механическое отсечение некоторых параметров не всегда приведет к недискриминации (посредством обработки больших данных можно извлечь даже ту информацию, которая кандидатами не предоставлялась)²².

Вышеназванным Регламентом № 2018/1725 установлена обязательность оценки воздействия на защиту данных (data protection impact assessment). Если тип обработки данных, в частности использование новых технологий с учетом характера, масштаба, контекста и целей обработки, может привести к возникновению высокой степени риска для прав и свобод физических лиц, контролер перед обработкой должен провести оценку воздействия предполагаемых операций по обработке на защиту персональных данных (ст. 39). Для оценки соблюдения основных прав могут использоваться большие данные, которые содержат потенциал выявления систематических предубеждений и дискриминации²³.

Углубленные исследования утверждают, что суждение о нейтральности технологий, алгоритмов в частности, несколько преувеличено. Алгоритм составляется

людьми, они отбирают его составные части, преследуя определенную цель. Это означает, что в алгоритм уже закладывается предпочтение определенной ориентации или поведения²⁴. Получается, что выбор был уже сделан на стадии разработки алгоритма.

Чтобы предоставить гражданам «защиту» от алгоритмов, во Франции, например, согласно Закону о цифровой республике от 7 октября 2016 г. правила, определяющие основные алгоритмические обработки, должны публиковаться на сайте государственного органа. Прозрачность алгоритмов становится своеобразной формой контроля, а прозрачность алгоритмического принятия решений служит тому, чтобы несправедливые дискриминации могли быть обнаружены и оспорены²⁵.

В принимаемом административном решении об использовании алгоритма должны отражаться способы функционирования алгоритма²⁶. Но как доступно и понятно объяснить алгоритм, если это математика, а не литература? Мало его расшифровать, информация должна быть доступна для человеческого понимания. Кроме того, прозрачность алгоритмов имеет и обратную сторону — когда человек использует полученное знание для удовлетворения личных интересов (например, чтобы сформулировать, подогнать свои данные под алгоритм с целью выигрыша конкурса).

²² См.: *Hacker Ph.* The Ambivalence of Algorithms. Gauging the Legitimacy of Personalized Law // *Bakhoun M., Conde Gallego B., Mackenrodt M.-O., Surblytė-Namavičienė G.* (eds.). *Personal Data in Competition, Consumer Protection and Intellectual Property Law. Towards a Holistic Approach?* Springer, 2018. P. 87. URL: <https://doi.org/10.1007/978-3-662-57646-5>.

²³ URL: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

²⁴ См.: *Pauliat H.* La décision administrative et les algorithmes: une loyauté à consacrer // *Revue du droit public*. 2018. No. 3. P. 643.

²⁵ См.: *Sartor G.* Human Rights and Information Technologies // *The Oxford Handbook of Law, Regulation and Technology* / ed. by *R. Brownsword, E. Scotford, K. Yeung*. Oxford, 2017. P. 442—450. DOI: 10.1093/oxfordhb/9780199680832.013.79.

²⁶ См.: *Duclercq J.-B.* L'automatisation algorithmique des décisions administratives individuelles // *Revue du droit public*. 2019. No. 2. P. 313.

Наконец, искусственный интеллект может ошибаться. При определении ответственности за ошибку, совершенную искусственным интеллектом, прежде следует ответить на вопрос: выполняет ли компьютер набор правил, которые приводят к предсказуемым результатам, или же компьютер способен перенастроить свои операции для достижения результатов, разрешающих проблемы непредвиденными способами?²⁷ Чем умнее система, тем менее она предсказуема. Поэтому если искусственный интеллект может менять алгоритм, правомерно поставить вопрос о разработке проблематики юридической личности. В любом случае алгоритмы, закладываемые в искусственный интеллект, должны предусматривать безусловную возможность их отключения человеком²⁸.

Применение алгоритмов и искусственного интеллекта в период пандемии коронавируса. Мировой опыт использования цифровых технологий во время пандемии подчеркнул хрупкость прав человека. В имеющемся опыте условно можно выделить азиатский и западный подходы.

В азиатских странах (Китай, Южная Корея, Тайвань, Сингапур) установка отслеживающих приложений на телефоны граждан была скорее обязательным, нежели добровольным решением (хотя и поддержанным большинством населения). И все же пользователи социальных сетей жаловались на отсутствие прозрачности в том, как работает приложение и какие данные оно хранит. Некоторые из них сообщили, что не могут изменить ошибочные «красные» обозначения

(цветные коды здоровья), и поставили под сомнение зависимость от подключения к Интернету и слежения²⁹.

Кстати, к этому блоку стран можно отнести также Израиль³⁰, где в середине марта 2020 г. было разрешено использовать методы массового электронного наблюдения (обычно применяемые для антитеррористической борьбы на палестинских территориях) для сдерживания эпидемии. Практически это означает, что внутренняя разведывательная служба Shin Bet с согласия генерального прокурора и без предварительного разрешения суда может отслеживать геолокацию мобильных телефонов зараженных лиц (в течение 30 дней), а также тех, кто находились возле них (в течение 14 дней). Анализ этих данных передается министерству здравоохранения, которое должно предупредить посредством смс-уведомления людей, рискующих быть зараженными, о необходимости помещения в карантин.

То, что удалось реализовать в Азии, безусловно привлекательно с точки зрения эффективности сдерживания пандемии, но порождает вопросы с позиции соблюдения прав человека и основ правового государства. Группой ученых из Оксфорда³¹ создано приложение для отслеживания контактов, которое стро-

²⁹ См.: China's coronavirus health code apps raise concerns over privacy. URL: <https://www.theguardian.com/world/2020/apr/01/chinas-coronavirus-health-code-apps-raise-concerns-over-privacy#maincontent> (дата обращения: 19.04.2020).

³⁰ URL: https://www.lemonde.fr/international/article/2020/03/17/israel-approuve-des-methodes-de-surveillance-electronique-de-masse-contre-le-coronavirus_6033390_3210.html#xtor=AL-32280270 (дата обращения: 04.05.2020).

³¹ URL: <https://science.sciencemag.org/content/early/2020/04/09/science.abb6936> (дата обращения: 04.05.2020).

²⁷ См.: Hildebrandt M. Criminal Law and Technology in a Data-Driven Society // Dubber M. D., Hörnle T. (eds.). The Oxford Handbook of Criminal Law. Oxford, 2014. P. 174—200. DOI: 10.1093/oxfordhdb/9780199673599.013.0009.

²⁸ См.: Малышкин А. В. Указ. соч. С. 456.

ится на запоминании близких контактов и немедленном уведомлении контактов о подтвержденных случаях COVID-19. Утверждается, что технология позволяет достичь эпидемического контроля при условии, что такое приложение используется достаточным количеством людей. Такая возможность сдерживать эпидемии без массовых карантинных порождает массу вопросов этического характера. В рамках работы оксфордской группы проводились опросы населения в ряде европейских стран. Любопытно, что две трети опрошиваемых выступили за автоматическую установку такого приложения, но при этом высказаны опасения, связанные: 1) с рисками кибератак на телефоны с установленным приложением; 2) возможностью продления наблюдения после окончания пандемии; 3) возрастанием психологического беспокойства в связи с использованием приложения; 4) техническими трудностями установки приложения.

Использование Bluetooth на смартфонах стало технологически безусловным фаворитом: если инфицированный COVID-19 контактировал с некими людьми, они незамедлительно об этом информируются и уведомляются о необходимости помещения в карантин. При этом проинформированные люди не знают, кто именно заражен и где именно они с ним контактировали. Такой подход, отвечающий политике сохранения конфиденциальности персональных данных, использован в США (PrivateKit), Австралии (CovidSafe), Италии (Immun), Франции (StopCovid) и др. При этом правительственное отношение к нему могло меняться по мере распространения эпидемии — так, еще в конце марта французское правительство заявляло, что подобное мобильное приложение противоречит французской культуре и нарушает фундаментальные права. Однако развитие эпидемии изменило эту позицию, и проект стал частью пра-

вительственного плана по выходу из режима самоизоляции. Он даже рассматривается как укрепление национального суверенитета, поскольку управляется только французскими компаниями, а не интернет-гигантами. Но нужно подчеркнуть два незыблемых условия для применения приложения — добровольность и анонимность.

В России предварительную оценку использованным во время пандемии средствам дал Совет при Президенте РФ по развитию гражданского общества и правам человека. В его докладе³² технологии, применяемые в целях обеспечения режима самоизоляции граждан, разделены на две группы в зависимости от субъектов воздействия:

1) технологии, обеспечивающие соблюдение режима карантина гражданами, инфицированными COVID-19 (обязательные к установке мобильные приложения, контроль геолокации мобильных устройств, отслеживание перемещений граждан и автотранспорта при помощи городских камер видеонаблюдения);

2) технологии, обеспечивающие соблюдение режима самоизоляции гражданами, не инфицированными COVID-19 (цифровые пропуска для личных нужд и передвижения автотранспорта, ограничение использования транспортных карт, регламентация прогулок вблизи места жительства).

Совет сделал вывод, что названные технологии обладают рядом преимуществ по сравнению с традиционными способами контроля (так, цифровой пропуск сложнее подделывать, а автоматизированный контроль обладает более высокой степенью выявления нарушений), однако и степень риска нарушения консти-

³² См.: Доклад Совета при Президенте РФ по развитию гражданского общества и правам человека по теме «Уроки эпидемии с точки зрения соблюдения прав и свобод человека и гражданина». URL: <http://president-sovet.ru/documents/read/687/>.

туционных прав граждан несоизмеримо выше.

Следует добавить и другие печальные девиации во время пандемии — незаконный сбор данных, их ошибочная трактовка (известен случай, когда штраф за нарушение режима карантина автоматически выписывался множество раз, когда человек просто перемещался по квартире: она находилась на границе Москвы и Московской области и хождение в ней квалифицировалось как перемещение между регионами).

В общем, к технике вопросов много. Но еще больше — к людям.

Проведенный анализ убеждает в том, что алгоритмы и искусственный интеллект найдут применение

в публично-правовых отношениях. Такое применение должно основываться на глубоком теоретическом осмыслении, технических и социальных расчетах, на учете имеющегося опыта, в том числе ошибочного. Очевидно, что от алгоритмов, примененных искусственным интеллектом, способны пострадать фундаментальные права человека — равенство, право на частную жизнь и защиту персональных данных, право на судебную защиту против решений, принятых алгоритмом. Эти вопросы, особенно потенциальная дискриминационность критериев обработки данных, нуждаются в детальной проработке и последующем нормативном регулировании.

Библиографический список

Boudinar-Zabaleta A. Algorithmes et lignes directrices. Réflexions sur la codification automatisée des motifs des décisions administratives // *Droit administratif*. 2019. No. 4.

Bourcier D. De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? // *Droit et société*. 2001. No. 3.

China's coronavirus health code apps raise concerns over privacy. URL: <https://www.theguardian.com/world/2020/apr/01/chinas-coronavirus-health-code-apps-raise-concerns-over-privacy#maincontent> (дата обращения: 19.04.2020).

Duclercq J.-B. L'automatisation algorithmique des décisions administratives individuelles // *Revue du droit public*. 2019. No. 2. P. 313.

Duclercq J.-B. Le droit public à l'ère des algorithmes // *Revue du droit public*. 2017. No. 5.

Hacker Ph. The Ambivalence of Algorithms. Gauging the Legitimacy of Personalized Law // Bakhom M., Conde Gallego B., Mackenrodt M.-O., Surblytė-Namavičienė G. (eds.). *Personal Data in Competition, Consumer Protection and Intellectual Property Law. Towards a Holistic Approach?* Springer, 2018. URL: <https://doi.org/10.1007/978-3-662-57646-5>.

Hildebrandt M. Criminal Law and Technology in a Data-Driven Society // Dubber M. D., Hörnle T. (eds.). *The Oxford Handbook of Criminal Law*. Oxford, 2014. DOI: 10.1093/oxfordhb/9780199673599.013.0009.

Huret A., Huet J.-M. L'intelligence artificielle au service du marketing // *L'Expansion Management Review*. 2012. No. 3.

Kamiran F., Žliobaitė I., Calders T. Quantifying explainable discrimination and removing illegal discrimination in automated decision making // *Knowledge and Information Systems*. 2013. No. 35. DOI: 10.1007/s10115-012-0584-8.

Pauliat H. La décision administrative et les algorithmes: une loyauté à consacrer // *Revue du droit public*. 2018. No. 3.

Sartor G. Human Rights and Information Technologies // *The Oxford Handbook of Law, Regulation and Technology* / ed. by R. Brownsword, E. Scotford, K. Yeung. Oxford, 2017. DOI: 10.1093/oxfordhb/9780199680832.013.79.

Бартенев Д. Г. Запрет дискриминации: практический обзор подходов Европейского суда по правам человека // *Международное правосудие*. 2019. № 1.

Близнец И. А., Леонтьев К. Б., Кубышкин А. В. Правовая охрана авторских и смежных прав в Российской Федерации. М., 2017.

Ворожевич А. С., Заиченко Е. В., Кирсанова Е. Е. и др. Современные информационные технологии и право: монография / отв. ред. Е. Б. Лаутс. М., 2019.

Доклад Совета при Президенте РФ по развитию гражданского общества и правам человека по теме «Уроки эпидемии с точки зрения соблюдения прав и свобод человека и гражданина». URL: <http://president-sovet.ru/documents/read/687/>.

Ковлер А. И. Права человека в цифровую эпоху // Бюллетень Европейского суда по правам человека. Российское издание. 2019. № 6.

Малышкин А. В. Интегрирование искусственного интеллекта в общественную жизнь: некоторые этические и правовые проблемы // Вестник Санкт-Петербургского университета. Право. 2019. № 3.

Пашенцев Д. А. Особенности правоприменения в условиях цифровизации общественных отношений // Вестник Санкт-Петербургского университета. Право. 2020. № 1.

Российское гражданское право: в 2 т. Т. 1: Общая часть. Вещное право. Наследственное право. Интеллектуальные права. Личные неимущественные права: учебник / отв. ред. Е. А. Суханов. М., 2011.

Савельев А. И. Научно-практический постатейный комментарий к Федеральному закону «О персональных данных». М., 2017.

Хабриева Т. Я., Черногор Н. Н. Право в условиях цифровой реальности // Журнал российского права. 2018. № 1.

Цивилистическая концепция интеллектуальной собственности в системе российского права: монография / под общ. ред. М. А. Рожковой. М., 2018.

Южаков В. Н., Талапина Э. В., Добролюбова Е. И., Тихомиров Ю. А. Инициативный проект закона об обеспечении качества государственного управления. М., 2020.

Algorithms and Artificial Intelligence in the Human Rights Context

E. V. Talapina

Institute of State and Law, Russian Academy of Sciences, Moscow 119019, Russian Federation

E-mail: comparative_law@igpran.ru

DOI: 10.12737/jrl.2020.118

The program acts on the artificial intelligence development address the prospects for its using in public management, which is a new scientific and practical problem. The law (especially private law) knows the algorithms, which are based on artificial intelligence as only objects of legal protection. For public law, the task of theoretical understanding of technological renewal is relevant, taking into account the knowing legal values such as fundamental human rights. By means of traditional methods of legal research, the author analyzes the Russian and foreign legal doctrine, judicial practice, examples from law enforcement practice in the field of approaches to regulation and use of technologies in public legal relations. This has highlighted the challenges that the human rights protection system may face.

The algorithm's idea as a set of instructions that can solve a problem and achieve a certain given result well meets the concept of good public administration, since the result of management can be embedded in the algorithm for making a management decision. The greatest efficiency of algorithms is predicted in areas where there is no discretionary power, which is characteristic of the provision of public services. But the management decisions automation carries possible risks.

In public law the scope of algorithms (or rather automatic data processing) is limited to personal data. It is the threat to personal data that is one of the risks in automatic decision-making. The technology of data processing consists in applying certain specified criteria to a particular person. A direct threat to fundamental human rights is contained in the potential discriminatory nature of these criteria. The right to non-discrimination is derived from the principle of equality. But formal legal equality also means the possibility of applying a general rule to a particular case, taking into account individual circumstances, whereas the algorithm impartially provides arithmetic equality. Moreover, artificial intelligence, trained on the basis of real and non-ideal human decisions, can perceive discrimination as a "normal" element of the algorithm. The identified problems need detailed studying and subsequent regulatory solution.

Keywords: algorithm, artificial intelligence, discrimination, personal data, equality, public services.

For citation: Talapina E. V. Algorithms and Artificial Intelligence in the Human Rights Context. *Zhurnal rossijskogo prava = Journal of Russian Law*, 2020, no. 10, pp. 25—39. DOI: 10.12737/jrl.2020.118 (In Russ.)

References

- Bartenev D. G. Prohibition of discrimination: a practical review of the approaches of the European Court of Human Rights. *Mezhdunarodnoe pravosudie*, 2019, no. 1, pp. 43—66. (In Russ.)
- Bliznets I. A., Leont'ev K. B., Kubyshkin A. V. *Pravovaya okhrana avtorskikh i smezhnykh prav v Rossiyskoy Federatsii*. Moscow, 2017. 260 p.
- Boudinar-Zabaleta A. Algorithmes et lignes directrices. Réflexions sur la codification automatisée des motifs des décisions administratives. *Droit administratif*, 2019, no. 4, pp. 13—18.
- Bourcier D. De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? *Droit et société*, 2001, no. 3, pp. 847—871.
- China's coronavirus health code apps raise concerns over privacy. Available at: <https://www.theguardian.com/world/2020/apr/01/chinas-coronavirus-health-code-apps-raise-concerns-over-privacy#maincontent> (accessed 19.04.2020).
- Doklad Soveta pri Prezidente RF po razvitiyu grazhdanskogo obshchestva i pravam cheloveka po teme "Uroki epidemii s tochki zreniya soblyudeniya prav i svobod cheloveka i grazhdanina". Available at: <http://president-sovet.ru/documents/read/687/>.
- Duclercq J.-B. L'automatisation algorithmique des décisions administratives individuelles. *Revue du droit public*, 2019, no. 2, pp. 295—320.
- Duclercq J.-B. Le droit public à l'ère des algorithmes. *Revue du droit public*, 2017, no. 5, pp. 1401—1433.
- Hacker Ph. The Ambivalence of Algorithms. Gauging the Legitimacy of Personalized Law. In Bakhom M., Conde Gallego B., Mackenrodt M.-O., Surblytė-Namavičienė G. (eds.). *Personal Data in Competition, Consumer Protection and Intellectual Property Law. Towards a Holistic Approach?* Springer, 2018. Pp. 85—117. Available at: <https://doi.org/10.1007/978-3-662-57646-5>.
- Hildebrandt M. Criminal Law and Technology in a Data-Driven Society. In Dubber M. D., Hörnle T. (eds.). *The Oxford Handbook of Criminal Law*. Oxford, 2014. Pp. 174—200. DOI: 10.1093/oxfordhb/9780199673599.013.0009.
- Huret A., Huet J.-M. L'intelligence artificielle au service du marketing. *L'Expansion Management Review*, 2012, no. 3, pp. 18—26.
- Kamiran F., Žliobaitė I., Calders T. Quantifying explainable discrimination and removing illegal discrimination in automated decision making. *Knowledge and Information Systems*, 2013, no. 35, pp. 613—644. DOI: 10.1007/s10115-012-0584-8.
- Khabrieva T. Y., Chernogor N. N. The Law in the Conditions of Digital Reality. *Zhurnal rossijskogo prava = Journal of Russian Law*, 2018, no. 1, pp. 85—102. DOI: 10.12737/art_2018_1_7 (In Russ.)
- Kovler A. I. Prava cheloveka v tsifrovuyu epokhu. *Byulleten' Evropeyskogo suda po pravam cheloveka. Rossiyskoe izdanie*, 2019, no. 6, pp. 146—150.
- Malyshkin A. V. Integration of artificial intelligence into public life: some ethical and legal problems. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2019, no. 3, pp. 444—460. (In Russ.)
- Pashentsev D. A. Features of law enforcement in the conditions of digitalization of social relations. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2020, no. 1, pp. 35—49. (In Russ.)
- Pauliat H. La décision administrative et les algorithmes: une loyauté à consacrer. *Revue du droit public*, 2018, no. 3, pp. 641—650.
- Rossiyskoe grazhdanskoe pravo. Vol. 1: Obshchaya chast'. Veshchnoe pravo. Nasledstvennoe pravo. Intellektual'nye prava. Lichnye neimushchestvennye prava. Ed. by E. A. Sukhanov. Moscow, 2011. 961 p.
- Sartor G. Human Rights and Information Technologies. *The Oxford Handbook of Law, Regulation and Technology*. Ed. by R. Brownsword, E. Scotford, K. Yeung. Oxford, 2017. Pp. 424—450. DOI: 10.1093/oxfordhb/9780199680832.013.79.

Savel'ev A. I. Nauchno-prakticheskiy postateynnyy kommentariy k Federal'nomu zakonu "O personal'nykh dannykh". Moscow, 2017. 318 p.

Tsivilisticheskaya kontseptsiya intellektual'noy sobstvennosti v sisteme rossiyskogo prava. Ed. by M. A. Rozhkova. Moscow, 2018. 269 p.

Vorozhevich A. S., Zaychenko E. V., Kirsanova E. E. et al. Sovremennye informatsionnye tekhnologii i pravo. Ed. by E. B. Lauts. Moscow, 2019. 288 p.

Yuzhakov V. N., Talapina E. V., Dobrolyubova E. I., Tikhomirov Yu. A. Initsiativnyy projekt zakona ob obespechenii kachestva gosudarstvennogo upravleniya. Moscow, 2020. 150 p.