МИР ВОКРУГ НАС

ИССЛЕДОВАТЕЛЬСКАЯ СТАТЬЯ

Искусственный интеллект в государственном управлении и транспарентность: европейский опыт

ИРИНА АНАТОЛЬЕВНА ЧЕРЕШНЕВА[a]
a Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации

DOI: 10.22394/2070-8378-2022-24-2-80-87

Аннотация: Статья посвящена правовым особенностям использования технологии искусственного интеллекта (далее – ИИ), новая волна интереса к которому вызвана бурной цифровой трансформацией. Зарождаясь и «расцветая» в частном секторе, она получает свое развитие и в сфере публичного интереса. В этой связи она нуждается в скрупулезном анализе, учитывающем, что как прямо, так и косвенно технология ИИ может становиться причиной нарушения прав человека. Автором отмечается, что объективная сторона транспарентности может быть выражена следующей триадой параметров: объяснимость, понятность, доступность; субъективная составляющая может быть представлена правом человека на объяснение полученного ИИ (алгоритмом) результата. В этом качестве право на объяснение должно быть отнесено к цифровым правам человека.

Сделан вывод, что при разработке и управлении технологиями искусственного интеллекта ни чрезмерное регулирование, ни полное невмешательство не являются подходящими или работоспособными. Предложено обеспечивать тестирование возможности и условий применения технологии искусственного интеллекта, в том числе в государственном управлении, в рамках экспериментального правового режима в сфере цифровых инноваций. С этой целью проведен анализ опыта обработки данных в государственном управлении Европейского союза, традиционно уделяющего повышенное внимание правам человека и защите данных. Статья подготовлена на основе научно-исследовательской работы, выполняемой в рамках государственного задания РАНХиГС.

Ключевые слова: транспарентность, государственное управление, искусственный интеллект, право на объяснение, Европейский союз, права человека, цифровизация
Дата поступления статьи в редакцию: 15 февраля 2022 года.

Artificial intelligence in public administration and its transparency: European experience

RESEARCH ARTICLE

IRINA ANATOLIEVNA CHERESHNEVA[a]
a Russian Presidential Academy of National Economy and Public Administration

Abstract: The article examines the legal features of artificial intelligence technology (hereinafter – AI) and the rapid digital transformation that causes the increased interest in this topic. Emerging and developing in the private sector, it is also taking shape in the sphere of public interest. In this regard, we need to conduct a thorough analysis of this technology, considering that, both directly and indirectly, AI technology can cause human rights violations. The author notes that the objective side of transparency can be expressed by the following triad of parameters: explainability, comprehensibility, and accessibility; the subjective component can be represented by the human right to an explanation of the obtained AI result (algorithm). To this extent, the right to an explanation should be classified as a digital human right.

The author concludes that neither excessive regulation nor complete non-interference is appropriate or functional in developing and managing AI technologies. The article proposes to facilitate further testing of the possibility and conditions of AI technology applications, including in public administration, as part of the experimental legal regime in the area of digital innovations. In this context, the author analyzed the experience of data processing in the public administration of the European Union, which has traditionally paid increased attention to human rights and data protection. The article is based on the research work carried out within the framework of the state assignment of RANEPA.

Keywords: transparency, public administration, artificial intelligence, right to explanation, European Union, human rights, digitalization
Received: February 15, 2022.

Введение

Технология искусственного интеллекта, так же как и другие виды прорывных цифровых технологий (например, технологии распределенного реестра, большие данные и др.), прежде всего нацелены на проведение в жизнь цифровой трансформации общества, глобальная задача которой заключается в обеспечении благосостояния, которое подразумевает не только улучшение социально-экономической обстановки, но и в некотором смысле освобождение человека от его повседневной рутины за счет автоматизации различного рода процессов. При этом подобные технологии, зарождаясь и расцветая в частном секторе (к примеру, технологические гиганты, финтех), мигрируют в сферу государственного. Несмотря на прорывной потенциал этих технологий, они нуждаются в скрупулезном анализе, поскольку как прямо, так и косвенно могут явиться причиной нарушения прав человека [Ковлер, 2019. С. 148].

В свою очередь, государственное управление в правовом государстве не может осуществляться иначе, нежели на основе права. Правовое осмысление искусственного интеллекта, несмотря на его формирующуюся историю и практические применения, находится на начальной стадии. Сказанное, однако, не указывает на отсутствие необходимости исследования существующих возможностей, способных минимизировать негативную сторону цифровых технологий (в частности, дискриминация, появление «общества слежки», ослабление человеческой автономии, искажение информации, вмешательство в выборы и цифровое отчуждение), одновременно высветив их положительные характеристики. В этой связи в качестве предмета исследования были избраны вопросы транспарентности искусственного интеллекта при его использовании в государственном управлении на примере опыта Европейского союза. Этот выбор закономерен ввиду того, что России близок европейский подход к обработке данных в государственном управлении, традиционно уделяющий повышенное внимание правам человека и защите данных.

Технология искусственного интеллекта: современное понимание

Повышению интереса к технологии искусственного интеллекта способствовало, в частности, возрастание возможностей обработки данных – доступность производственных мощностей, технологии больших данных и машинное обучение. Последнее выступает подразделом ИИ и относится к разработке цифровых систем, которые со временем улучшают свою производительность по ранее заданной задаче, обучаясь на собственном опыте. Функционирование алгоритма машинного обучения сравнивают с когнитивным развитием ребенка – правила воспроизводятся через наблюдение и анализ, без их точного объяснения [Villani et al., 2018. P. 26]. Кроме того, машинное обучение является наиболее широко используемой формой ИИ, с ним наиболее часто связывают машинное обучение, ведь именно оно внесло большой вклад в развитие беспилотных автомобилей, технологий распознавания речи и машинного перевода. Выделяются следующие виды машинного обучения:

– контролируемое обучение, которое позволяет модели ИИ обучаться на размеченных тренировочных данных (например, обучение модели ИИ помогать помечать контент на сайтах, используя теги);

– неконтролируемое обучение, которое обучает алгоритм ИИ использовать неразмеченную и неклассифицированную информацию;

– стимулирующее обучение, которое позволяет модели ИИ учиться в процессе выполнения задачи.

Таким образом, если рассматривать технологию искусственного интеллекта через призму машинного обучения, то она представляет собой: вовлечение вычислительных машин, использующих статистику для поиска закономерностей в больших объемах данных (1); способность выполнять повторяющиеся задачи с данными без необходимости наличия постоянного руководства со стороны человека (2)[1].

При этом отметим, что с нормативной точки зрения приходится констатировать отсутствие четкого единообразного определения ИИ. Так, в Национальной стратегии развития искусственного интеллекта на период до 2030 года[2] искусственный интеллект определяется 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. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе то, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений. Аналогичное определение дается в приказе Министерства экономического развития РФ от 29 июня 2021 года № 392 «Об утверждении критериев определения принадлежности проектов к проектам в сфере искусственного интеллекта».

В то же время в документах Европейского союза использован более широкий подход: искусственный интеллект – это системы, которые характеризует интеллектуальное поведение за счет анализа окружающей среды и принятия мер, с определенной степенью автономности и для достижения конкретных целей [L'intelligence…, 2018. P. 1]. Например, в Европейской этической хартии об использовании искусственного интеллекта в судебных системах и окружающих их реалиях искусственный интеллект понимается как комплекс науки, теории и техник,

1 A guide to using artificial intelligence in the public sector. https://www.gov.uk/government/publications/understanding-artificial-intelligence/a-guide-to-using-artificial-intelligence-in-the-public-sector

2 Указ Президента Российской Федерации от 10 октября 2019 года № 490 «О развитии искусственного интеллекта в Российской Федерации»

цель которых состоит в том, чтобы воспроизвести с помощью машины когнитивные возможности человека[3].

В свою очередь, в литературе выделяются следующие сущностные признаки ИИ: 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 обстоятельства [Малышкин, 2019. С. 446].

Транспарентность как неотъемлемое условие использования искусственного интеллекта в государственном управлении

В юридической доктрине применительно к государственному управлению под «транспарентностью» понимают «особый режим функционирования публичной власти, устанавливающий открытость отношений как между всеми субъектами, участвующими в управлении, так и в отношениях с обществом» [Талапина, 2021(a). С. 34]. В этом смысле транспарентность может быть представлена в качестве «двусторонней связи в контексте взаимодействия органов власти и граждан» [Кряжков, 1988. С. 14]. Содержание подобной связи может быть выражено как сочетание, с одной стороны, «предоставления государственными органами информации о своей деятельности самостоятельно и по запросу граждан, а с другой – как предоставление гражданам и организациям возможности выразить свою точку зрения по значимым общественным вопросам» [Силкин, 2021. С. 24]. Представляется, что и в отношении транспарентности использования ИИ в государственном управлении ключевой будет задача установления и поддержания взаимодействия (в сущности, связи) между государством и пользователем.

Как отмечается в зарубежной литературе, первостепенная важность транспарентности (и предполагаемой, и фактической) систем искусственного интеллекта, используемых в государственном управлении, обусловливается необходимостью обеспечения объяснимости как процесса работы искусственного интеллекта, так и полученного результата широкой общественности (а не только экспертам), поскольку открытость алгоритма позволяет понять, как и почему было принято то или иное управленческое решение [The use of Artificial Intelligence…, 2020. P. 16]. Отметим, что в отечественной Национальной стратегии развития искусственного интеллекта на период до 2030 года прозрачность (транспарентность) относится к одному из основных принципов развития и использования технологий искусственного интеллекта. Она понимается как объяснимость работы искусственного интеллекта и процесса достижения им результатов, недискриминационный доступ пользователей продуктов, которые созданы с использованием технологий искусственного интеллекта, к информации о применяемых в этих продуктах алгоритмах работы искусственного интеллекта.

Кроме того, подлинная транспарентность тесным образом связана с подотчетностью или ответственностью в значении восприимчивости государственных органов к запросам о предоставлении информации о каком-либо процессе или готовности предложить обоснование уже предпринятых или только планируемых действий. Оправдание этих ожиданий лежит в основе правомерного использования ИИ в государственном управлении.

На системном уровне это отражает необходимость достижения тонкого баланса между полной транспарентностью, с одной стороны, и достаточно комплексной и сложноорганизованной деятельностью в сфере государственного управления – с другой.

На техническом уровне речь идет о рассмотрении ИИ в качестве «черного ящика». «Черный ящик» – это описание, применяемое к некоторым системам глубокого обучения, которые принимают входной сигнал и обеспечивают выходной, но вычисления, которые происходят между этими процессами, нелегко интерпретировать людям. Системы ИИ типа «черный ящик», часто основанные на машинном обучении на больших данных, принимают решения экспериментально или интуитивно, в отсутствие возможности раскрытия причин, почему было принято то или иное решение [Талапина, 2021(b). С. 18]. В государственном управлении использование ИИ типа «черный ящик» может вызывать определенные сложности, обусловленные не только отсутствием транспарентности, но и потенциальной предвзятостью, встроенной в алгоритм и основанной на человеческих предубеждениях (а также программными компонентами сбора), скрытых в обучающих данных, которые могут привести к несправедливым или неверным решениям.

В противовес системам ИИ типа «черный ящик» существует так называемый «объяснимый искусственный интеллект» (Explainable AI – XAI)[4], в рамках которого результаты принятого системой ИИ решения могут быть поняты людьми. В этой связи представляется, что в сфере государственного управления преимущество должен получить именно объяснимый ИИ; однако приходится констатировать, что данная область исследования находится на этапе становления (даже по сравнению с «классическим» ИИ типа «черный ящик»).

При этом подчеркнем, что транспарентность использования ИИ включает в себя не только объяснимость как процесса его работы, так и полученного результата, но также их понятность (пользователь в состоянии усвоить, осмыслить, постигнуть содержание этого процесса, включить его в систему устоявшихся идей и представлений) [Ивин, Никифоров, 1997. С. 270]. Кроме того, нельзя не отметить тесную связь между транспарентностью ИИ и правом на объяснение (right to explanation) [Goodman, Flaxman, 2017. P. 52], под которым понимается право человека на получение объяснения относительно полу-

3 Европейская этическая хартия об использовании искусственного интеллекта в судебных системах и окружающих их реалиях. https://rm.coe.int/ru-ethical-charter-en-version-17-12-2018-mdl-06092019-2-/16809860f4

4 Explainable artificial intelligence. https://en.wikipedia.org/w/index.php?title=Explainable_artificial_intelligence&oldid=1069461374

ченного алгоритмом результата. Это становится особенно актуальным, когда речь заходит о принятых в отношении физического лица решениях, имеющих юридический или финансовый характер.

Так, если человек получает отказ в отношении поданной им заявки на получение кредита, то он может затребовать от кредитной организации информацию (объяснение), разъясняющую ему основания, в соответствии с которыми было принято решение об отказе [Berry et al., 2021. P. 229].

На этическом уровне речь идет не только об ответственности за причинение вреда, но и о взаимосвязи транспарентности и этического использования ИИ [The use of Artificial Intelligence…, 2020. P. 16]. Так, важным документом, определяющим основные принципы функционирования искусственного интеллекта на уровне Совета Европы, стала принятая 4 декабря 2018 года Европейская этическая хартия использования искусственного интеллекта в судебных системах и окружающих их реалиях[5]. Отметим, что принцип прозрачности, беспристрастности и достоверности выступает одним из принципов использования ИИ в судебной и правоохранительной системах.

Кроме того, транспарентность также означает доступность. Содержательные объяснения алгоритма вполне возможны, но они могут быть недоступны. Право интеллектуальной собственности может препятствовать раскрытию проприетарного кода или доступу к обучающим данным, поэтому, даже если и можно было бы понять, как работает алгоритм, полные вычисления могут оказаться невозможными по экономическим, юридическим или политическим причинам.

Создание и поддержание общественного доверия и уверенности в использовании ИИ требует обеспечения для обычного пользователя определенного уровня объяснения и понимания того, как работает ИИ, как он пришел к тому или иному результату, а также общедоступности такого объяснения. Соответственно, «способность объяснять» должна выступать важной частью процесса отбора / проектирования на этапе разработки систем ИИ. Несмотря на то, что существуют объективные причины, по которым подробное изложение функционирования систем ИИ может оказаться невыполнимым (например, соблюдение закона, право интеллектуальной собственности, техническая сложность), неспособность предложить даже базовое объяснение работы алгоритма существенно влияет на общественное доверие к ИИ [The use, 2020. P. 18]. В отсутствие доверия к технологии граждане не будут принимать участия в процессе принятия государственных решений, что, в свою очередь, подрывает легитимность и эффективность государственного сектора в целом.

Таким образом, применительно к использованию ИИ в государственном управлении транспарентность должна быть представлена следующей триадой ее параметров: объяснимость, понятность, доступность.

Нормативные требования к транспарентности ИИ: европейский опыт

В Европе сфера применения автоматизированной обработки данных ограничена персональными данными [Талапина, 2020. С. 28]. Так, ст. 5 Общего Регламента ЕС (далее – Регламент ЕС)[6] закрепляет принципы, связанные с обработкой персональных данных. Один из них гласит, что персональные данные должны обрабатываться законно, беспристрастно и прозрачным образом в отношении субъекта данных. Тем самым, устанавливается принцип транспарентности (прозрачности) обработки персональных данных. При этом системное толкование положений Регламента ЕС позволяет раскрыть содержание принципа транспарентности обработки персональных данных: любая информация и сведения субъекту данных должны предоставляться в сжатой, прозрачной, понятной и легкодоступной форме на понятном и простом языке (ст. 14 Регламента ЕС).

С учетом обозначенной выше связи транспарентности использования ИИ в государственном управлении и права на объяснение целесообразно кратко остановиться на рассмотрении вопроса о нормативном закреплении права на объяснение в Регламенте ЕС.

С одной стороны, в отношении автоматизированной обработки (к которой можно отнести и обработку данных с использованием ИИ) субъект данных должен иметь право на то, чтобы не подпадать под действие решения, которое может включать в себя меру для оценки относящихся к нему персональных аспектов, основанное исключительно на автоматизированной обработке и порождающее юридические последствия в отношении субъекта данных или аналогичным образом существенно влияющее на него (например, автоматический отказ в заявке на получение кредита онлайн). Помимо этого, указанная обработка должна осуществляться в соответствии с надлежащими гарантиями, включающими в себя особое информирование субъекта данных и право на вмешательство оператора, на высказывание своей точки зрения, получение объяснения в отношении решения, принятого после оценки, а также на оспаривание такого решения (параграф 71 Регламента ЕС).

С другой стороны, право на объяснение присутствует в необязательной части Регламента ЕС, в то время как в положениях, содержащих нормы права, оно не упоминается (было исключено в ходе законопроектной деятельности) [Edwards, Veale, 2017. P. 21]. Круг примене-

5 Европейская этическая хартия об использовании искусственного интеллекта в судебных системах и окружающих их реалиях. https://rm.coe.int/ru-ethical-charter-en-version-17-12-2018-mdl-06092019-2-/16809860f4

6 Регламент Европейского парламента и Совета Европейского союза 2016/679 от 27 апреля 2016 года о защите физических лиц при обработке персональных данных и о свободном обращении таких данных, а также об отмене Директивы 95/46/ЕС (Общий Регламент о защите персональных данных / General Data Protection Regulation / GDPR). https://mobileonline.garant.ru/#/document/71936226/paragraph/305/doclist/7622/showentries/0/highlight/General%20Data%20Protection%20Regulation%20|GDPR:2

ния данного права достаточно узок и охватывает только те решения, которые исключительно основываются на автоматизированной обработке и которые порождают юридические или аналогичные существенные последствия. Тем самым существует вероятность неприменения данного права в отношении тех случаев, которые могут представлять публичный интерес [Wachter et al., 2017. P. 82].

Сторонники выделения самостоятельного права на объяснение, относя его к социальным правам, рассматривают данное право в качестве важнейшей основы информационного общества, ввиду того, что институты современного общества все чаще прибегают (в будущем необходимость в «цифре» только усилится) к использованию различных цифровых технологий, включая искусственный интеллект [Berry et al., 2021. P. 230]. Другими словами, если в соответствующих автоматизированных системах принятия решений изначально заложить их объяснимость, то они не только получат статус «заслуживающие доверия», но и станут более транспарентными. Без этого права, которое бы могло получить свое закрепление как в нормативных (например, в законах о защите персональных данных), так и в ненормативных правовых актах (например, в профессиональных стандартах), общественность рискует получить минимальный набор специальных средств, в том числе правовых, направленных на оспаривание решений автоматизированных систем. Однако в этой связи возникает проблема: каким образом и в какой форме (текст, визуальная диаграмма высокого уровня, видео и др.) следует донести объяснение процесса и результата работы алгоритма до пользователя?

Противники данного права указывают на отсутствие необходимости в его выделении или рассматривают его в качестве сдерживающего фактора, препятствующего развитию технологий. Критика сводится к следующему: 1) предпочтение отдается решению, принятому человеком, а не машиной; 2) избыточность существующего регулирования[7]; 3) потенциальная невозможность объяснить пользователю некоторые алгоритмы машинного обучения.

Нам представляется, что в сфере государственного управления при использовании ИИ окончательное решение всегда должно оставаться за человеком. В этой связи подчеркнем, что важно убедиться в том, что участие человека действительно имеет смысл, а также признать факт того, что оно может различаться в зависимости от ситуации.

Сказанное означает, что государственные органы должны периодически проверять полученные ИИ результаты, соотнося их с определенными метриками, или параллельно выполнять ту же задачу силами человека и сопоставлять друг с другом результаты, полученные человеком и ИИ. Значимость человеческого участия может быть учтена посредством формирования междисциплинарных групп для разработки систем ИИ, в состав которых должны входить ученые-программисты, специалисты по выработке государственной политики и юристы. Такие группы могли бы иметь больше возможностей для обеспечения того, чтобы результаты использования ИИ не были бы недостаточно продуманными или несоответствующими политическим и законодательным намерениям.

В свою очередь, тезис об избыточности регулирования представляется нам несостоятельным. В защиту своей позиции приведем мнение Специального комитета по искусственному интеллекту (CAHAI), который призван содействовать разработке правовых рамок использования искусственного интеллекта в Европе[8]; им была проведена оценка целесообразности регулирования ИИ, которая выявила существенный пробел в правовом регулировании вопросов, связанных с проектированием, разработкой и использованием ИИ[9].

На наш взгляд, единственным заслуживающим внимания аргументом является тезис о том, что многие алгоритмы машинного обучения нелегко объяснить. Например, результат глубокой нейронной сети зависит от множества уровней вычислений, связанных сложным образом, и ни один ввод или вычисление не может выступать в качестве доминирующего фактора [Miller, 2019. P. 22]. Однако здесь на помощь может прийти объяснимый ИИ.

Кроме того, первостепенное значение в контексте правового регулирования искусственного интеллекта приобретают вопросы, связанные с разработкой конкретных требований к ИИ, которые бы обеспечили транспарентность, качество тренировочных данных, человеческую автономию, достаточные возможности для подачи жалобы и реализации средств правовой защиты, а также гарантировали эффективные механизмы правоприменения, поэтому необходимо:

1) надлежащее международное регулирование искусственного интеллекта, основанное на существующих правовых стандартах, но также дополняющее их, поскольку это имеет принципиальное значение для обеспечения соответствия проектирования, разработки и использования ИИ правам человека, демократии и верховенству права и их дальнейшему укреплению;

2) разработка обязательных горизонтальных правовых рамок, закладывающих базовые основы для любого применения ИИ, в том числе в государственном управлении;

3) разработка специальных отраслевых правовых документов обязательного или необязательного характера с учетом рисков, связанных с нарушением прав человека (дискриминация, профилирование, слежка (на-

7 EU's Right to Explanation: A Harmful Restriction on Artificial Intelligence. https://www.techzone360.com/topics/techzone/articles/2017/01/25/429101-eus-right-explanation-harmful-restriction-artificial-intelligence.htm

8 CAHAI – Ad hoc Committee on Artificial Intelligence. https://www.coe.int/en/web/artificial-intelligence/cahai

9 Conference Human Rights in the Era of AI. https://www.coe.int/en/web/artificial-intelligence/human-rights-in-the-era-of-ai

блюдение), ослабление человеческой автономии, цифровое отчуждение и др.)[10].

В этой связи нам представляется, что в контексте обеспечения транспарентности использования ИИ в государственном управлении право на объяснение должно получить свое нормативное закрепление на межгосударственном уровне, впоследствии став органической частью национальных правопорядков, признающих права и свободы человека высшей ценностью. С учетом стремительного развития цифровых технологий, где персональные данные рассматриваются в качестве «информационной нефти», необходимость в защите цифровых прав человека (к которым, на наш взгляд, можно отнести и право на объяснение) не подлежит сомнению. Отсюда видятся закономерными попытки разработки правовых рамок, в границах которых технология искусственного интеллекта сможет получить свое развитие, продолжая балансировать на грани прорыва и вызова.

Так, в апреле 2021 года был опубликован проект Регламента ЕС о гармонизированных правилах в отношении ИИ (Artificial Intelligence Act) (далее – проект Европейского регламента по ИИ)[11]. В целом документ выстроен на базовой идее соблюдения и эффективного применения существующего законодательства об основных правах человека.

В рамках регулирования использования систем искусственного интеллекта в Европе предлагается полностью запретить те из них, которые могут применяться не по назначению, предоставляя новые и мощные инструменты для манипулирования, эксплуатации и социального контроля. Использование систем ИИ не по назначению особенно вредно и должно быть запрещено, поскольку это противоречит таким ценностям ЕС, как уважение человеческого достоинства, свобода, равенство, демократия и верховенство права, и подрывает основные права, включая право на недискриминацию, право на защиту и неприкосновенность частной жизни и права детей.

Отметим, что особые правила устанавливаются проектом Европейского регламента по ИИ в отношении систем искусственного интеллекта, которые создают высокий риск для здоровья и безопасности, а также основных прав человека. Например, закрепляется, что системы ИИ высокого риска должны быть спроектированы и разработаны таким образом, чтобы обеспечить достаточную транспарентность их функционирования, которая позволит пользователям интерпретировать результаты работы системы и использовать их надлежащим образом. К ним также должны прилагаться инструкции по использованию в соответствующем цифровом формате или иным образом, включающие краткую, полную, точную и конкретную информацию, являющуюся релевантной, доступной и понятной для пользователей. При этом такая информация должна определять:

(a) личность и контактные данные провайдера;

(b) конкретные характеристики, возможности и ограничения производительности системы ИИ высокого риска (в частности, предполагаемая область применения; определенный уровень точности, устойчивости и кибербезопасности, в рамках которого данная система была протестирована и прошла валидацию; производительность системы ИИ и др.);

(c) определенные изменения в системе ИИ высокого риска и ее производительность, которые были предопределены провайдером на момент первоначальной оценки соответствия, в случае их наличия;

(d) меры по обеспечению надзора со стороны человека, направленные на облегчение интерпретации пользователями результатов систем ИИ;

(e) предполагаемый срок службы данной системы, а также необходимое техническое обслуживание и поддержка для обеспечения ее надлежащего функционирования (ст. 13 проекта Европейского регламента по искусственному интеллекту).

Выводы

На основании вышеизложенного подведем итоги настоящего исследования, а также постараемся наметить направления для дальнейшей работы.

Во-первых, базовым принципом использования ИИ в государственном управлении является транспарентность. С одной стороны, она означает объяснимость (процесса работы ИИ и полученного им результата), с другой – включает общедоступность такого объяснения для широкой общественности. В этой связи необходимо не только установить презумпцию транспарентности использования искусственного интеллекта в государственном управлении, но и по умолчанию закрепить обязанность раскрытия государственными органами целей использования ИИ (за исключением тех сфер государственного управления (например, оборона и безопасность), где такое раскрытие может угрожать публичным интересам). В данном ракурсе можно говорить об объективной стороне транспарентности использования ИИ в государственном управлении.

В свою очередь, субъективная сторона транспарентности может быть представлена правом на объяснение, под которым понимается право человека на получение объяснения относительно полученного алгоритмом результата. Это право, на наш взгляд, может быть отнесено к цифровым правам человека, ввиду чего оно должно получить свое нормативное закрепление, прежде всего на межгосударственном уровне, а в последующем – в законодательстве о защите персональных данных. Кроме того, дополнительной гарантией реализации права на объяснение может выступить предложение об установлении нормативным правовым актом порядка, в соответствии с которым будет исключено использование оценки, проведенной ИИ, в качестве единственной основы для принятия решения должностным лицом, обладающим

10 Conference Human Rights in the Era of AI. https://www.coe.int/en/web/artificial-intelligence/human-rights-in-the-era-of-ai

11 Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

соответствующей компетенцией и наделенным соответствующими дискреционными полномочиями. Ведь, как справедливо отмечается в литературе, «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 оценки» [Пашенцев, 2020. С. 42].

Во-вторых, транспарентность использования ИИ в государственном управлении предполагает наличие открытого программного кода (свободное программное обеспечение, то есть право пользователя свободно запускать, копировать, распространять, изучать, изменять и улучшать его[12]) в противовес частному сектору, где главенствует его проприентарный собрат (программное обеспечение, принадлежащее правообладателю (субъект частного права) на праве интеллектуальной собственности). В свою очередь, проприетарность алгоритма может явиться не только причиной нарушения основных прав и свобод человека, но также выступить препятствием для реализации принципа транспарентности ИИ[13]. В этой связи необходимо достижение баланса, с одной стороны, между транспарентностью и правом интеллектуальной собственности, ввиду того, что мы не можем исключить ситуацию, при которой наработки частного сектора (в том числе разработанный программный код) могут найти свое применение в сфере государственного управления; с другой – между транспарентностью и действительной возможностью ее реализации, обусловленной сложной технологической составляющей систем ИИ.

В первом случае можно предусмотреть частичную публикацию ключевой информации, касающейся разработки, проектирования и эксплуатации систем ИИ: об использованных алгоритмах, целях, преследуемых при оптимизации алгоритмов, подборе данных для настройки программы, средних значениях и типовых отклонениях в полученных результатах или же об объеме и типе данных, обрабатываемых алгоритмом[14]. Во втором – можно использовать простой и понятный язык для описания способа достижения результатов (посредством передачи информации о характере предлагаемых услуг, разработанных инструментов, положительных результатах и рисках ошибок). Кроме того, независимые органы, в том числе по защите данных, могли бы обеспечить сертификацию, предварительный и/или последующий контроль методов обработки. В свою очередь, контрольно-надзорная деятельность органов государственной власти могла бы получить свою реализацию в форме выдачи свидетельств[15].

В-третьих, при разработке и управлении технологиями ИИ, в том числе в государственном управлении, ни чрезмерное регулирование, ни подход, основанный на полном невмешательстве, не являются подходящими или работоспособными. В этой связи целесообразно обеспечить тестирование ИИ [Ефремов и др., 2020. С. 16], например, в рамках Федерального закона от 31 июля 2020 года № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в РФ», Федерального закона от 24 апреля 2020 года № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте РФ – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных»». Арсенал экспериментального правового режима в сфере цифровых инноваций позволит протестировать использование ИИ в сфере государственного управления, в том числе и реализацию принципа транспарентности. Ведь задача права как раз и заключается в том, чтобы учесть технические наработки и снабдить их арсеналом традиционных средств реагирования.

В заключение отметим, что транспарентность – это важнейший публичный аспект качественного государственного управления (обоснованность, результативность, эффективность) [Добролюбова и др., 2019. С. 10], а в контексте цифровизации и цифровой трансформации государственного управления обеспечение транспарентности использования технологии искусственного интеллекта приобретает критически важное значение. В противном случае мы рискуем столкнуться с ситуацией, при которой произвольное использование алгоритмов (ИИ) неминуемо приведет к серьезным нарушениям прав человека, и баланс между прорывом и вызовом будет утерян.

12 Свободное программное обеспечение. https://ru.bmstu.wiki/index.php?title=Free_software&mobileaction=toggle_view_mobile

13 Подробнее см.: *Черешнева И.А.* Алгоритмы и право: анализ, вдохновленный судебным прецедентом. Информационное общество. 2021. № 6. С. 84–90.

14 Европейская этическая хартия об использовании искусственного интеллекта в судебных системах и окружающих их реалиях. https://rm.coe.int/ru-ethical-charter-en-version-17-12-2018-mdl-06092019-2-/16809860f4

15 Европейская этическая хартия об использовании искусственного интеллекта в судебных системах и окружающих их реалиях. https://rm.coe.int/ru-ethical-charter-en-version-17-12-2018-mdl-06092019-2-/16809860f4

Литература

Добролюбова Е.И., Южаков В.Н., Ефремов А.А., Клочкова Е.Н., Талапина Э.В., Старцев Я.Ю. Цифровое будущее государственного управления по результатам. М.: Издательский дом «Дело» РАНХиГС, 2019.

Ефремов А.А., Добролюбова Е.И., Талапина Э.В., Южаков В.Н. Экспериментальные правовые режимы: зарубежный опыт и российский старт. М.: Издательский дом «Дело» РАНХиГС, 2020.

Ивин А.А., Никифоров А.Л. Словарь по логике. М., 1997.

Ковлер А.И. Права человека в цифровую эпоху. *Бюллетень Европейского суда по правам человека. Российское издание*. 2019. № 6. С. 146–150.

Кряжков В.А. Информационные связи советской представительной системы в условиях демократизации общества: автореф. дис. ... д-ра юрид. наук. Свердловск, 1988.

Малышкин А.В. Интегрирование искусственного интеллекта в общественную жизнь: некоторые этические и правовые проблемы. *Вестник Санкт-Петербургского университета. Право*. 2019. № 3. С. 444–460.

Пашенцев Д.А. Особенности правоприменения в условиях цифровизации общественных отношений. *Вестник Санкт-Петербургского университета. Право.* 2020. № 1. С. 35–49.

Силкин В.В. Транспарентность исполнительной власти в цифровую эпоху. *Российский юридический журнал*. 2021. № 4. С. 20–31.

Талапина Э.В. Алгоритмы и искусственный интеллект сквозь призму прав человека. *Журнал российского права*. 2020. № 10. С. 25–39.

Талапина Э.В. Использование искусственного интеллекта в государственном управлении. *Информационное общество*. 2021(a). № 3. С. 16–22.

Талапина Э.В. Перспективы правового регулирования качества публичного управления. *Административное право и процесс*. 2021(b). № 11. С. 33–36.

Berry D.M., Balaskas B. (ed.) Explanatory Publics: Explainability and Democratic Thought // Fabricating Publics: The Dissemination of Culture in the Post-truth Era. Open Humanities Press. 2021. P. 211–233. In English

Edwards L., Veale M. Slave to the Algorithm? Why a Right to Explanationn is Probably Not the Remedy You are Looking for. *SSRN Electronic Journal*. 2017. Vol. 16. P. 1–67. In English

Goodman B., Flaxman S. European Union regulations on algorithmic decision-making and a «right to explanation». *AI Magazine*. 2017. Vol. 38. № 3. P. 50–57. In English

L'intelligence artificielle pour l'Europe. Communication de la Commission Européenne. Bruxelles, 2018.

Miller T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence*. 2019. Vol. 267. P. 1–38. In English

The use of Artificial Intelligence to Combat Public Sector Fraud. Professional Guidance. International Public Sector Fraud Forum. 2020. In English

Villani C., Schoenauer M., Bonnet Y., Berthet Ch., Cornut A.-Ch., Levin F., Rondepierre B. Donner un sens à l'intelligence artificielle: Pour une stratégie nationale et européenne. Mission Villani sur l'intelligence artificielle. 2018.

Wachter S., Mittelstadt B., Floridi L. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*. 2017. Vol. 7. Issue 2. P. 76–99. In English

References

Dobrolyubova E.I., Yuzhakov V.N., Efremov A.A., Klochkova E.N., Talapina E.V., Startsev Ya.Yu. The digital future of public administration by results. Moscow: Izdatel'skiy dom "Delo" RANKhiGS, 2019. In Russian

Efremov A.A., Dobrolyubova E.I., Talapina E.V., Yuzhakov V.N. Experimental legal regimes foreign experience and Russian start. Moscow: Izdatel'skiy dom "Delo" RANKhiGS, 2020. In Russian

Ivin A.A., Nikiforov A.L. Dictionary of logic. Moscow: Gumanitarnyi izdatel'skiy tsentr VLADOS, 1997. In Russian

Kovler A.I. Human rights in the digital age. *Byulleten' Yevropeyskogo suda po pravam cheloveka. Rossiyskoye izdaniye.* 2019. No. 6. P. 146–150. In Russian

Kryazhkov V.A. Information relations of the Soviet representative system in the context of the society's democratization: abstract of the dissertation of a Doctor of Legal Sciences. Sverdlovsk, 1988. In Russian

Malyshkin A.V. Integrating artificial intelligence into public life: some ethical and legal problems. *Vestnik Sankt-Peterburgskogo universiteta. Pravo.* 2019. No. 3. P. 444–460. In Russian

Pashentsev D.A. Features of law enforcement in the context of digitalization of public relations. *Vestnik Sankt-Peterburgskogo universiteta. Pravo.* 2020. No. 1. P. 35–49. In Russian

Silkin V.V. Transparency of executive power in digital era. *Rossiyskiy yuridicheskiy zhurnal*. 2021. No. 4. P. 20–31. In Russian

Talapina E.V. Algorithms and artificial intelligence in the context of human rights. *Zhurnal rossiyskogo prava.* 2020. No. 10. P. 25–39. In Russian

Talapina E.V. The use of artificial intelligence in public administration. *Informatsionnoye obshchestvo.* 2021(a). No. 3. P. 16–22. In Russian

Talapina E.V. Prospects of the legal regulation of the quality of public administration. *Administrativnoye pravo i protsess.* 2021(b). No. 11. P. 33–36. In Russian

Artificial intelligence for Europe. Communication from the European Commission. Brussels. 2018. In French

Villani C., Bonnet Y., Schoenauer M., Berthet Ch., Levin F., Cornut A.Ch., Rondepierre B. For a meaningful artificial intelligence: towards a French and European strategy. Villani mission on artificial intelligence. 2018. In French

Информация об авторе:

ИРИНА АНАТОЛЬЕВНА ЧЕРЕШНЕВА, младший научный сотрудник Центра технологий государственного управления Института прикладных экономических исследований Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации (Российская Федерация, 119571, Москва, проспект Вернадского, 82). E-mail: chereshneva-ia@ranepa.ru

Ссылка для цитирования: *Черешнева И.А.* Искусственный интеллект в государственном управлении и транспарентность: европейский опыт. *Государственная служба*. 2022. № 2. С. 80–87.

Information about the author:

IRINA ANATOLIEVNA CHERESHNEVA, Junior Researcher, Center for Public Administration Technologies, Institute of Applied Economic Research Russian Presidential Academy of National Economy and Public Administration (82, Vernadsky Prospekt, Moscow, 119571, Russian Federation). E-mail: chereshneva-ia@ranepa.ru

For citation: *Chereshneva I.A.* Artificial intelligence in public administration and its transparency: European experience. *Gosudarstvennaya sluzhba*. 2022. No. 2. P. 80–87.