

Д. Л. Кутейников\*,  
О. А. Ижаев\*\*,  
С. С. Зенин\*\*\*,  
В. А. Лебедев\*\*\*\*

## Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика»<sup>1</sup>

**Аннотация.** В статье изучены правовые подходы Европы и Соединенных Штатов Америки, на которых основывается законодательство, регулирующее использование компьютерных алгоритмов, т.е. систем автоматизированного принятия юридически значимых решений. Установлено, что в указанных юрисдикциях применяются существенно разные концепции.

Европейский подход предусматривает регулирование применения систем автоматизированного принятия решений посредством законодательства о персональных данных. Авторы приходят к выводу о том, что Общий регламент по защите данных не возлагает юридической обязанности на контролеров раскрывать техническую информацию, т.е. открывать «черный ящик», субъекту персональных данных, в отношении которого алгоритм принимает решение. Это может случиться в перспективе, когда законодатель-

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-29-16193.

© Кутейников Д. Л., Ижаев О. А., Зенин С. С., Лебедев В. А., 2020

\* *Кутейников Дмитрий Леонидович*, кандидат юридических наук, старший преподаватель кафедры конституционного и муниципального права Московского государственного юридического университета имени О.Е. Кутафина (МГЮА)  
Садовая-Кудринская ул., д. 9, г. Москва, Россия, 125993  
kuteynikov@me.com

\*\* *Ижаев Осман Аликович*, кандидат юридических наук, консультант юридического отдела Московской городской избирательной комиссии  
Моховая ул., д. 11, стр. 8, г. Москва, Россия, 125009  
izhaev.osman@gmail.com

\*\*\* *Зенин Сергей Сергеевич*, кандидат юридических наук, директор Научно-исследовательского института, доцент кафедры конституционного и муниципального права Московского государственного юридического университета имени О.Е. Кутафина (МГЮА), ведущий научный сотрудник кафедры теории государства и права, конституционного и административного права Южно-Уральского государственного университета (национального исследовательского университета)  
Садовая-Кудринская ул., д. 9, г. Москва, Россия, 125993  
zeninsergei@mail.ru

\*\*\*\* *Лебедев Валериан Алексеевич*, доктор юридических наук, профессор, профессор кафедры конституционного и муниципального права Московского государственного юридического университета имени О.Е. Кутафина (МГЮА)  
Садовая-Кудринская ул., д. 9, г. Москва, Россия, 125993  
lebedev\_va@bk.ru

ные органы конкретизируют положения названного Регламента, в соответствии с которыми контролер должен предоставлять субъекту персональных данных значимую информацию о логике принимаемых в его отношении решений.

В США вопросы прозрачности и подотчетности алгоритмов регламентируются при помощи различных антидискриминационных актов, регулирующих отдельные сферы жизнедеятельности человека. При этом они носят фрагментарный характер и их совокупность не представляет собой комплексную, связанную между собой систему нормативных правовых актов. На практике правовое регулирование осуществляется ad hoc со ссылкой на отдельные законодательные положения, запрещающие обработку чувствительных видов персональных данных.

В статье утверждается, что в России правовое регулирование алгоритмической прозрачности и подотчетности находится на этапе становления. Имеющееся законодательство о персональных данных позволяет предположить, что отечественный подход к разрешению проблемы «черного ящика» близок к европейскому. При разработке и принятии соответствующих нормативных правовых актов необходимо исходить из того, что субъект персональных данных должен обладать правом на получение информации, объясняющей в доступной форме логику принятого в его отношении решения.

**Ключевые слова:** алгоритмы; машинное обучение; искусственный интеллект; большие данные; алгоритмическая подотчетность; алгоритмическая прозрачность; «черный ящик»; персональные данные; приватность.

**Для цитирования:** Кутейников Д. Л., Ижаев О. А., Зенин С. С., Лебедев В. А. Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика» // Lex russica. — 2020. — Т. 73. — № 6. — С. 139—148. — DOI: 10.17803/1729-5920.2020.163.6.139-148.

## Algorithmic Transparency and Accountability: Legal Approaches to Solving the "Black Box" Problem<sup>2</sup>

**Dmitriy L. Kuteynikov**, Cand. Sci. (Law), Senior Lecturer of the Department of Constitutional and Municipal Law, Kutafin Moscow State Law University (MSAL)  
ul. Sadovaya-Kudrinskaya, d. 9, Moscow, Russia, 125993  
kuteynikov@me.com

**Osman A. Izhaev**, Cand. Sci. (Law), Consultant of the Legal Department of the Moscow City Election Commission  
ul. Mokhovaya, d. 11, str. 8, Moscow, Russia, 125009  
izhaev.osman@gmail.com

**Sergey S. Zenin**, Cand. Sci. (Law), Head of the Research Institute, Associate Professor of the Department of Constitutional and Municipal Law, Kutafin Moscow State Law University (MSAL), Leading Researcher of the Department of Theory of the State and Law, Constitutional and Administrative Law of the South Ural State University (National Research University)  
ul. Sadovaya-Kudrinskaya, d. 9, Moscow, Russia, 125993  
zeninsergei@mail.ru

**Valerian A. Lebedev**, Dr. Sci. (Law), Professor, Professor of the Department of Constitutional and Municipal Law, Kutafin Moscow State Law University (MSAL)  
ul. Sadovaya-Kudrinskaya, d. 9, Moscow, Russia, 125993  
lebedev\_va@bk.ru

**Abstract.** The paper examines the European and American legal approaches based on legislation regulating the use of computer algorithms, i.e. systems for automated decision-making of legally significant decisions. It is established that these jurisdictions apply essentially different concepts.

The European approach provides for regulating the use of automated decision-making systems through legislation on personal data. The authors conclude that the general data protection regulation does not impose

<sup>2</sup> The reported study was funded by RFBR according to the research project № 18-29-16193.

a legal obligation on the controllers to disclose technical information, i.e. to open a "black box", to the subject of personal data, in respect of which the algorithm makes a decision. This may happen in the future, when the legislative authorities specify the provisions of this Regulation, according to which the controller must provide the subject of personal data with meaningful information about the logic of decisions taken in relation to it.

In the United States, issues of transparency and accountability of algorithms are regulated by various anti-discrimination acts that regulate certain areas of human activity. At the same time, they are fragmentary and their totality does not represent a complex, interconnected system of regulatory legal acts. In practice, legal regulation is carried out ad hoc with reference to certain legal provisions prohibiting the processing of sensitive types of personal data.

The paper states that the legal regulation of algorithmic transparency and accountability is in its infancy in Russia. The existing legislation on personal data suggests that the domestic approach to solving the "black box" problem is close to the European one. When developing and adopting relevant regulatory legal acts, it is necessary to proceed from the fact that the subject of personal data should have the right to receive information explaining the logic of the decision made in relation to it in an accessible form.

**Keywords:** algorithms; machine learning; artificial intelligence; big data; algorithmic accountability; algorithmic transparency; "black box"; personal data; privacy.

**Cite as:** Kuteynikov DL, Izhaev OA, Zenin SS, Lebedev VA. Algoritmicheskaya prozrachnost i podotchetnost: pravovye podkhody k razresheniyu problemy «chernogo yashchika» [Algorithmic Transparency and Accountability: Legal Approaches to Solving the "Black Box" Problem]. *Lex russica*. 2020;73(6):139-148. DOI: 10.17803/1729-5920.2020.163.6.139-148. (In Russ., abstract in Eng.).

## Введение

В современном мире наблюдается устойчивая тенденция развития и массового внедрения компьютерных алгоритмов. Они способны воспринимать окружающую среду, анализировать полученную из нее информацию и на этой основе принимать рациональные решения с разной степенью автономности. Особенностью данных алгоритмов является то, что в некоторых случаях невозможно с высокой степенью достоверности восстановить ход принятия им того или иного решения<sup>3</sup>. Таким образом, функционирование алгоритмов является непрозрачным и неподотчетным, что в литературе принято называть проблемой «черного ящика».

Данная проблема приобретает особое для права значение в случае использования таких

алгоритмов в общественной сфере. Принимаемые ими решения могут иметь юридическое значение или являться основой для принятия человеком решений, порождающих правовые последствия. При таких условиях субъекты соответствующих правоотношений не имеют возможности оценить обоснованность и справедливость принятых алгоритмом решений и тем самым де-факто лишаются права на их обжалование. Между тем за этой тайной принятия решений могут скрываться высокие риски нарушения прав личности.

К настоящему времени накоплен достаточный опыт применения алгоритмов. В частности, широко известны случаи их отрицательного воздействия на гражданские<sup>4</sup>, политические<sup>5</sup>, экономические<sup>6</sup> и социальные<sup>7</sup> права и свободы человека. В мировой

<sup>3</sup> Burrell J. How the machine «thinks». Understanding opacity in machine learning algorithms // *Big Data & Society*. 2016. № 1-12 (January — June).

<sup>4</sup> В некоторых штатах США судьи в ходе рассмотрения дел использовали алгоритм, который определял степень риска совершения обвиняемыми повторных преступлений. Независимая проверка алгоритма показала, что он присваивал лицам с темной кожей заведомо более высокую степень риска (почти в два раза) по сравнению с другими. Интересно, что коммерческая организация, которая разработала и применяла данный алгоритм, не пожелала раскрывать информацию относительно логики присваивания программным обеспечением той или иной степени риска. См.: Machine Bias There's software used across the country to predict future criminals. And it's biased against black // ProPublica. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (дата обращения: 30.03.2020).

<sup>5</sup> Известен печальный опыт компании Cambridge Analytica, которая при помощи алгоритма анализировала данные пользователей Facebook и на основании полученных результатов влияла через агрессивную таргетированную рекламу на политические предпочтения граждан, в частности в ходе избиратель-

практике существуют различные правовые подходы, направленные на избежание негативных последствий массового использования алгоритмов. В связи с этим представляет научный и практический интерес исследование двух качественно отличающихся друг от друга концепций правового регулирования прозрачности и подотчетности алгоритмов.

### **I. Европейский подход к раскрытию информации о логике работы систем принятия автоматизированных решений**

В Европейском Союзе функционирование алгоритмов регулирует Общий регламент по защите данных (далее — Регламент)<sup>8</sup>. Требования Регламента, предъявляемые к контролерам, по раскрытию информации в целях объяснения субъектам данных логики принятого алгоритмом в их отношении решения исследованы в ряде научных работ<sup>9</sup>. Непосредственно к вопросам прозрачности и подотчетности алгоритмов относятся положения Регламента, в соответствии с которыми контролер в случаях получения персональных данных от субъекта данных, а также от третьей

стороны должен предоставлять, а субъект вправе получать информацию «о наличии системы автоматизированного принятия решения, включая профилирование, указанное в пунктах 1 и 4 статьи 22 Регламента, и по крайней мере в этих случаях полноценную информацию о логике работы этой системы, а также о значимости и предполагаемых последствиях такой обработки для субъекта данных» (ст. 13—15 Регламента). При этом пункт 1 ст. 22 Регламента устанавливает, что «субъект данных имеет право не быть субъектом такого решения, которое основано исключительно на автоматической обработке, в том числе профилировании, которое в отношении субъекта данных создает правовые последствия или которое подобным образом существенно воздействует на субъект данных».

Вышеизложенные нормативные положения ученые интерпретируют по-разному.

Одни утверждают, что из этих положений вытекает обязанность контролеров при использовании алгоритмов, принимающих юридически значимые решения, раскрывать субъекту данных полную информацию о логике принятия решения, в том числе техническую, что, по сути, означает раскрытие «черного ящика»<sup>10</sup>.

ной кампании по выборам президента США в 2016 г. См.: *The Cambridge Analytica Files* // Guardian. URL: <https://www.theguardian.com/news/series/cambridge-analytica-files> (дата обращения: 30.03.2020).

<sup>6</sup> Все более популярным становится так называемое персонализированное ценообразование. Многие интернет-компании используют различные переменные для определения цены на товар или услугу, к примеру геолокацию покупателя, принадлежность к каким-либо меньшинствам. См.: *Uber's New Pricing Idea Is Good Theory, Risky Business* // Perma. URL: <https://perma.cc/2T8E-J5K8> (дата обращения: 30.03.2020); *Amazon Doesn't Consider the Race of Its Customers. Should It?* // Bloomberg. URL: <https://www.bloomberg.com/graphics/2016-amazon-same-day/> (дата обращения: 30.03.2020).

<sup>7</sup> Исследователи из Университета Карнеги — Мэллона, изучив алгоритм таргетированной рекламы Google, обнаружили, что пользователям Интернета мужского пола в шесть раз чаще, чем пользователям женского пола, показывали рекламу высокооплачиваемой работы. См.: *Datta A., Tschantz M. C., Datta A. Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination* // *Proceedings on Privacy Enhancing Technologies*. 2015. № 1. Pp. 92—112.

<sup>8</sup> *General Data Protection Regulation (GDPR)* // URL: <http://www.gdpr.eu/doc/gdpr%20reglment.odt> (дата обращения: 30.03.2020).

<sup>9</sup> *Malgieri G., Comandé G. Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation* // *International Data Privacy Law*. 2017. Vol. 7, Iss. 4. Pp. 243—265; *Selbst A. D., Powles J. Meaningful information and the right to explanation* // *International Data Privacy Law*. 2017. Vol. 7, Iss. 4. Pp. 233—242; *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, Iss. 2. Pp. 76—99; *Isak M., Lee A. B. The Right Not to Be Subject to Automated Decisions Based on Profiling* // *EU Internet Law: Regulation and Enforcement* / ed. T. Synodinou, P. Jougoux, C. Markou, T. Prastitou. Springer, 2017. Pp. 77—98; *Tae Wan Kim, Routledge B. Algorithmic Transparency, a Right to Explanation, and Placing Trust* // Carnegie Mellon University. URL: <https://static1.squarespace.com/static/592ee286d482e908d35b8494/t/59552415579fb30c014cd06c/1498752022120/Algorithmic+transparency%2C+a+right+to+explanation+and+trust+%28TWK%26BR%29.pdf> (дата обращения: 30.03.2020).

<sup>10</sup> *Malgieri G., Comandé G. Op. cit.*; *Selbst A. D., Powles J. Op. cit.*

Такой подход предполагает расширительное толкование норм Регламента.

Другие считают, что контролеры не обязаны предоставлять подробную информацию, а должны лишь уведомить субъекта персональных данных (не *ex-post*, а *ex-ante*) о том, как в общих чертах (абстрактно) алгоритм принимает решение<sup>11</sup>. Основание такого ограничительного толкования заключается в том, что Регламент не конкретизирует то, какую именно информацию необходимо предоставить субъекту персональных данных. Кроме того, важным аргументом здесь является то, что в ст. 22 Регламента имеется уточнение, что положения применяются в случаях, когда система приняла решение исключительно самостоятельно, т.е. без всякого вмешательства человека. На практике человек часто принимает формальное участие в принятии алгоритмом решения, что позволяет контролеру избегать соблюдения соответствующих положений Регламента.

С одной стороны, из практики имплементации в национальное законодательство данных положений странами — членами Европейского Союза следует, что подавляющее их большинство не стали использовать расширительный подход, лишь небольшая часть стран приняла дополнительные акты, обязывающие предоставлять субъектам данных подробную информацию о функционировании алгоритмов<sup>12</sup>. С другой стороны, есть также позиция Рабочей группы по защите прав человека в сфере обработки персональных данных, функционирующей под эгидой Европейской комиссии. В соответствии с ней из формулировки «решения,

основанные исключительно на автоматизированной обработке», не следует, что в случае, если человек принимает формальное участие, контролер освобождается от обязанности предоставлять соответствующую информацию субъектам персональных данных<sup>13</sup>.

Представляется, что из буквального прочтения положений Регламента действительно не вытекает наличие у контролеров обязанности раскрывать подробную информацию об алгоритме. Вместе с тем данный вопрос остается открытым и будет разрешен в дальнейшем правоприменительной практикой.

Несмотря на то что действующие на уровне Европейского Союза нормативные положения все еще не регулируют подробно данные вопросы, в научно-практической сфере наблюдается тенденция поиска сбалансированной правовой модели, которая будет содержать все необходимые юридические требования к контролерам и предоставлять достаточный перечень правовых средств защиты субъектам персональных данных, т.е. лицам, в отношении которых алгоритм будет принимать юридически значимые решения. В частности, об этом свидетельствует значительное количество научных исследований, осуществляемых под эгидой органов Европейского Союза и Совета Европы<sup>14</sup>.

В одном из таких исследований изучены социальные, технические и нормативные проблемы, связанные с алгоритмической прозрачностью и подотчетностью. В докладе, опубликованном по его результатам, отмечается, что целесообразно говорить о прозрачности в семи сферах машинного обучения:

<sup>11</sup> Wachter S., Mittelstadt B., Floridi L. Op. cit. ; Isak M., Lee A. B. Op. cit. ; Tae Wan Kim, Routledge B. Op. cit.

<sup>12</sup> Malgieri G. Automated decision-making in the EU Member States: The right to explanation and other «suitable safeguards» in the national legislations // Computer law & security review. 2019. № 35. P. 6.

<sup>13</sup> Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01) // European Commission. URL: [https://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=612053](https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053) (дата обращения: 30.03.2020).

<sup>14</sup> См., например: White Paper on Artificial Intelligence — A European approach to excellence and trust (2020) // European Commission. URL: [https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf) (дата обращения: 30.03.2020) ; A governance framework for algorithmic accountability and transparency (2019) // European Parliament. URL: [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS\\_STU\(2019\)624262\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf) (дата обращения: 30.03.2020) ; Understanding algorithmic decision-making: Opportunities and challenges (2019) // European Parliament. URL: [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS\\_STU\(2019\)624261\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf) (дата обращения: 30.03.2020) ; Addressing the impacts of Algorithms on Human Rights (2018) // Council of Europe. URL: <https://rm.coe.int/draft-recommendation-of-the-committee-of-ministers-to-states-on-the-hu/168095eecf> (дата обращения: 30.03.2020) ; Discrimination, artificial intelligence, and algorithmic decision-making (2018) // Council of Europe. URL: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> (дата обращения: 30.03.2020).

- 1) данные (раскрытие источника данных, методов, с помощью которых они были проверены как непредвзятые и репрезентативные, информация о процессах самостоятельного обновления и обучения);
- 2) непосредственно алгоритм (раскрытие сведений о разработке программного обеспечения, кода и порядка его работы);
- 3) цели (прозрачность в отношении задач и приоритетов алгоритма);
- 4) результаты (требование от производителей или операторов открытости информации об эксплуатации алгоритмических систем и последствий);
- 5) комплаенс (регулярная проверка алгоритма самими операторами и информирование о ее результатах);
- 6) влияние (сведения о наличии субъектов, в интересах которых используется алгоритм);
- 7) использование (информация о том, какие именно персональные данные и в каком объеме использует алгоритм).

При этом авторы утверждают, что прозрачность может иметь отношение либо в целом к системе, либо к принятию отдельных решений. Они приходят к выводу, что требовать открытости всего алгоритма нецелесообразно, поскольку это очень ресурсоемкий процесс и не ведет к значимому объяснению логики принятых решений. Авторы настаивают на том, что более разумным способом является установление мер, направленных на раскрытие сведений о логике принятия индивидуальных решений. В качестве таких мер предлагается анализ входных данных, статистическое объяснение, проверка конструкции/кода и статистический анализ, определение чувствительности отдельных данных (какие именно переменные определяют результат)<sup>15</sup>.

Таким образом, для европейского подхода характерна нацеленность на раскрытие информации относительно функционирования алгоритма, с тем чтобы субъекты имели полноценную возможность получить исчерпывающее обоснование принятого в их отношении решения. Вместе с тем данный подход не обязательно ведет к раскрытию «черного ящика», он базируется на тезисе, что в случае реальной необходимости и целесообразности, в отсутствие

альтернативных способов объяснения логики принятия решений возможно установление требований по раскрытию технических особенностей (кода) соответствующего алгоритма.

## II. Алгоритмическая подотчетность и прозрачность в Соединенных Штатах Америки

В Соединенных Штатах Америки на данный момент на федеральном уровне отсутствует закон, который регламентировал бы вопросы, связанные с прозрачностью и подотчетностью алгоритмов. Это обусловлено многочисленными факторами, в том числе особенностями правовой системы США.

Европейский Регламент по защите данных является сквозным нормативным правовым актом и потому затрагивает вопросы алгоритмической прозрачности и подотчетности в разных сферах. В США, напротив, оборот персональных данных регламентируется концептуально иным образом. В Европе применяется известное правило «всё, что прямо не разрешено, то запрещено», в то время как в США руководствуются принципом «всё, что прямо не запрещено, разрешено». Отдельные нормативные правовые положения, применение которых становится возможным в спорах, касающихся использования систем автоматизированного принятия решений, содержатся в не связанных между собой различных антидискриминационных законах на уровне федерации и штатов.

К примеру, федеральный закон Fair Housing Act<sup>16</sup> требует от кредиторов в специальном уведомлении заявителя на получение кредита указывать факторы, ведущие к отказу в его выдаче и иным негативным последствиям. Equal Credit Opportunity Act<sup>17</sup> предписывает, что кредиторы обеспечивают «изложение причин» отрицательных решений. В соответствии с аналогичными актами запрещается учитывать при принятии решений характеристики, обработка которых находится под запретом (например, пол, раса). Вместе с тем данный подход разрабатывался для предотвращения дискриминации, вытекающей из решений, принимаемых человеком, и потому не предоставляет исчерпывающего набора средств правовой защиты

<sup>15</sup> A governance framework for algorithmic accountability and transparency (2019).

<sup>16</sup> Pub L No 90-284, 82 Stat 81 (1968), codified as amended at 42 USC § 3601 et seq.

<sup>17</sup> Pub L No 94-239, 90 Stat 251 (1976), codified as amended at 15 USC § 1691 et seq.

от дискриминации со стороны систем автоматизированного принятия решений<sup>18</sup>.

Одной из попыток урегулировать данный вопрос на федеральном уровне является внесение в 2019 г. в Конгресс США законопроекта «Об алгоритмической подотчетности»<sup>19</sup>. Этот акт предусматривает обязанность ряда лиц, использующих в своей деятельности системы автоматизированного принятия решений, проводить оценку их воздействия в ходе ее разработки и использования, в том числе данные проектирования и обучения, на соответствие требованиям точности, справедливости, непредвзятости, недискриминационности, конфиденциальности и безопасности. Кроме того, в надзорный орган необходимо предоставлять подробное описание системы автоматизированного принятия решений, ее конструкции, обучения, данных и ее назначения. В то же время неизвестно, что имеется в виду под подробным описанием. Результаты проведенной оценки публикуются по усмотрению оператора.

Сходный законопроект предлагается принять в штате Калифорния<sup>20</sup>. В законопроекте содержится понятие системы автоматизированного принятия решений, представляющей собой вычислительный процесс, в том числе процесс, основанный на машинном обучении, статистике или других методах обработки данных, техниках искусственного интеллекта, который способен принимать решения или облегчает принятие человеком решений, воздействующих на людей. Наряду с этим предлагается ввести термин «оценка воздействия автоматизированной системы принятия решений», что означает

исследование, оценивающее систему автоматизированного принятия решений и процесс ее разработки, включая, но не ограничиваясь данными о дизайне и об обучении системы, на предмет влияния на точность, справедливость, предвзятость, недискриминационность, конфиденциальность и безопасность.

До 1 марта 2022 г. предлагается также создать консультативную целевую группу по системам автоматизированного принятия решений с целью рассмотрения и предоставления рекомендаций по их использованию в бизнесе, правительстве и других сферах.

В штатах Нью-Йорк<sup>21</sup> и Вашингтон<sup>22</sup> разрабатывается законодательство в сфере применения систем автоматизированного принятия решений в государственном секторе (государственные решения и судопроизводство).

### III. Проблема «черного ящика»: состояние правового регулирования в Российской Федерации

В ряде программных документов, посвященных развитию искусственного интеллекта в Российской Федерации, непрозрачность и неподотчетность алгоритмов упоминается в качестве проблемы, требующей своего нормативного правового разрешения<sup>23</sup>. Вместе с тем действующее российское законодательство фактически не содержит нормативных положений, регулирующих в достаточной мере вопросы прозрачности и подотчетности алгоритмов. Это обстоятельство не позволяет говорить о сформировавшемся оте-

<sup>18</sup> См., например: *Talia B. G., Jann S.* Big Data and Discrimination // *University of Chicago Law Review*. 2019. № 459. P. 462; *Bruckner M. A.* Promise and Perils of Algorithmic Lenders' Use of Big Data // *Chicago-Kent Law Review*. 2019. Vol. 93. Iss. 1. Pp. 56—57.

<sup>19</sup> H.R.2231 — Algorithmic Accountability Act of 2019 // *Congress.gov*. URL: <https://www.congress.gov/bill/116th-congress/house-bill/2231/text> (дата обращения: 30.03.2020).

<sup>20</sup> AB-2269 Personal rights: automated decision systems (2019—2020) // *California Legislative Information*. URL: [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB2269&search\\_keywords=algorithm](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB2269&search_keywords=algorithm) (дата обращения: 30.03.2020).

<sup>21</sup> Automated decision systems used by agencies // *New York City Council*. URL: <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0> (дата обращения: 30.03.2020).

<sup>22</sup> Senate Bill 5527 // *Lawfilesex*. URL: <http://lawfilesex.leg.wa.gov/biennium/2019-20/Pdf/Bills/Senate%20Bills/5527.pdf?q=20200215114844> (дата обращения: 30.03.2020).

<sup>23</sup> Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 года) // Президент России. URL: <http://www.kremlin.ru/acts/bank/44731> (дата обращения: 02.04.2020); Концепция регулирования технологий искусственного интеллекта и робототехники до 2023 года // Сколково. URL: <https://sk.ru/foundation/legal/m/sklegal11/23859/download.aspx> (дата обращения: 02.04.2020).

чественном подходе к разрешению проблемы «черного ящика».

Однако некоторые нормы, имеющие отношение к системам автоматизированного принятия решений, установлены в законодательстве, регулирующем оборот персональных данных. Так, в соответствии со ст. 16 Федерального закона «О персональных данных»<sup>24</sup> в Российской Федерации запрещается принимать решения, порождающие правовые последствия, на основании исключительно автоматизированной обработки персональных данных. Эта статья сформулирована таким образом, что, несмотря на общий запрет, разрешается принятие соответствующих решений, во-первых, при наличии согласия субъекта персональных данных, выраженного в письменной форме, во-вторых, в случаях, предусмотренных федеральными законами. Иными словами, законодательство о персональных данных не содержит исчерпывающего перечня случаев, при которых разрешается принимать решения на основании автоматизированной обработки персональных данных.

На наш взгляд, правовое регулирование использования алгоритмов должно оперировать термином, однозначно определяющим технологию принятия юридически значимых решений. В Федеральном законе «О персональных данных» таким понятием является «автоматизированная обработка персональных данных», что предполагает обработку персональных данных с помощью средств вычислительной техники. Этот термин, когда он используется применительно к принятию решений посредством вычислительных средств, имеет ряд недостатков. Во-первых, он определяет не саму технологию, а процесс, т.е. означает только определенную деятельность. Алгоритмы же представляют собой не процесс, а технологию, применение которой и требует правового регулирования. С этой точки зрения целесообразнее использовать термин, не только описывающий процесс, но и конкретизирующий саму технологию. Во-вторых, названный термин касается обработки только персональных данных, в то время как системы принятия решений анализируют в том числе и иные данные, не имеющие статуса персональных. В-третьих, сама по себе обработка данных не предполагает принятия конечного решения, такое понимание способно ввести

в заблуждение, поскольку может показаться, что решение принимает сам человек.

Таким образом, понятие «автоматизированная обработка персональных данных» при его использовании в контексте принятия юридически значимых решений не является релевантным и не отражает степени проникновения в общественную жизнь алгоритмов. Представляется, что для создания правовых условий, обеспечивающих подотчетность и прозрачность алгоритмов в Российской Федерации, следует использовать термин «система автоматизированного принятия решений». Это понятие лишено вышеописанных недостатков, является достаточно объемным и в то же время однозначным. Кроме того, оно доказало свою эффективность в зарубежных юрисдикциях, рассмотренных в настоящей статье.

Относительно прозрачности алгоритмов в Федеральном законе «О персональных данных» указано, что оператор обязан разъяснить субъекту персональных данных порядок принятия решения на основании автоматизированной обработки персональных данных, возможные юридические последствия, а также предоставить возможность заявить возражение против такого решения.

Однако разъяснение субъекту персональных данных порядка принятия решения на основании автоматизированной обработки персональных данных не означает, что субъекту персональных данных должна быть предоставлена информация относительно логики принятого решения. Рассматриваемые положения касаются не того, как именно принято решение, к примеру что положено в его основу, а, скорее, процедуры его принятия. Из указанного следует, что субъект персональных данных, несмотря на то что наделен правом подать возражение на решение, фактически им воспользоваться не сможет, поскольку не будет знать, как именно было принято решение. В то же время основополагающей задачей правового регулирования должно являться создание условий, при которых человек имеет право на получение исчерпывающей информации относительно логики принятого в его отношении решения. Достижение указанной цели не обязательно ведет к раскрытию «черного ящика», речь идет о реализации мер, направленных на

<sup>24</sup> Федеральный закон от 27.07.2006 № 152-ФЗ (ред. от 31.12.2017) «О персональных данных» // Официальный интернет-портал правовой информации. URL: <http://www.pravo.gov.ru>. 31.12.2017.

объяснение человеку причинно-следственной связи соответствующего решения.

Таким образом, можно утверждать, что в настоящее время законодательство, разрешающее проблему «черного ящика», фактически отсутствует. Вместе с тем в Российской Федерации существует законодательство о персональных данных, имеющее сквозной характер и распространяющееся на все сферы общественных отношений. С этой точки зрения для нашего государства, по нашему мнению, будет более целесообразно применять рассмотренный в настоящей статье европейский подход.

### Заключение

На уровне Европейского Союза нормативные требования к алгоритмической прозрачности и подотчетности инкорпорированы в законодательство, регулирующее оборот персональных данных. На контролеров возлагается обязанность предоставлять субъекту персональных данных определенный объем информации, связанной с процессом принятия алгоритмом в его отношении решения.

Европейский правовой подход основывается на примате прав человека над интересами субъектов, получающих экономическую и иную выгоду от использования алгоритмов. Вместе с тем такой подход еще не реализован в полной мере, поскольку отсутствует конкретизация положений Регламента по защите данных, касающихся работы алгоритмов. Так, из буквального понимания действующих положений Регламента следует, что они не возлагают на контролеров юридическую обязанность раскрывать техническую информацию (по сути, открывать «черный ящик») о работе алгоритма субъекту персональных данных.

Ключевой особенностью правового подхода Соединенных Штатов Америки является то,

что у них отсутствует сквозной нормативный правовой акт, устанавливающий какие-либо требования к алгоритмической прозрачности и подотчетности. На практике правовое регулирование осуществляется ad hoc со ссылкой на отдельные положения различных антидискриминационных актов, запрещающих обработку чувствительных видов персональных данных.

В то же время в США на уровне федерации и штатов парламентарии выступают с законодательными инициативами, в которых предлагается установление «обязательной оценки воздействия системы автоматизированного принятия решений», которая ведет к раскрытию определенного перечня сведений относительно работы алгоритма, а также регламентируются права и обязанности субъектов соответствующих правоотношений.

Относительно состояния правового регулирования в Российской Федерации вопросов алгоритмической прозрачности и подотчетности следует сказать, что оно находится на этапе становления. Имеющееся законодательство о персональных данных позволяет предположить, что отечественный подход близок к европейскому, которому присуще разрешение проблемы «черного ящика» через регулирование оборота персональных данных. В центре такого подхода находится право субъекта персональных данных на получение информации относительно логики принятого в его отношении решения. Разработка соответствующей нормативной правовой базы, которая обеспечит гарантии реализации этого права, является актуальной и необходимой задачей нашего государства. В противном случае существенно возрастут риски нарушения прав и свобод человека в эпоху массового внедрения в общественную практику технологий четвертой промышленной революции.

### БИБЛИОГРАФИЯ

1. *Bruckner M. A.* Promise and Perils of Algorithmic Lenders' Use of Big Data // Chicago-Kent Law Abstract. — 2019. — Vol. 93. — Iss. 1.
2. *Burrell J.* How the machine «thinks». Understanding opacity in machine learning algorithms // Big Data & Society. — 2016. — № 1-12 (January — June).
3. *Isak M., Lee A. B.* The Right Not to Be Subject to Automated Decisions Based on Profiling // EU Internet Law: Regulation and Enforcement / ed. T. Synodinou, P. Jougleux, C. Markou, T. Prastitou. — Springer, 2017.
4. *Malgieri G.* Automated decision-making in the EU Member States: The right to explanation and other «suitable safeguards» in the national legislations // Computer law & security review. — 2019. — № 35.

5. *Malgieri G., Comandé G.* Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation // *International Data Privacy Law*. — 2017. — Vol. 7, Iss. 4.
6. *Selbst A. D., Powles J.* Meaningful information and the right to explanation // *International Data Privacy Law*. — 2017. — Vol. 7, Iss. 4.
7. *Talia B. G., Jann S.* Big Data and Discrimination // *University of Chicago Law Abstract*. — 2019. — № 459.
8. *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, Iss. 2.

*Материал поступил в редакцию 17 апреля 2020 г.*

#### REFERENCES

1. Bruckner MA. Promise and Perils of Algorithmic Lenders' Use of Big Data. *Chicago-Kent Law Abstract*. 2019;93(1). (In Eng.)
2. Burrell J. How the machine «thinks». Understanding opacity in machine learning algorithms. *Big Data & Society*. 2016;1-12. (In Eng.)
3. Isak M, Lee AB. The Right Not to Be Subject to Automated Decisions Based on Profiling. In: Synodinou T, Jogleux P, Markou C, Prastitou T, editors. *EU Internet Law: Regulation and Enforcement*. Springer; 2017. (In Eng.)
4. Malgieri G. Automated decision-making in the EU Member States: The right to explanation and other «suitable safeguards» in the national legislations. *Computer law & security review*. 2019;35. (In Eng.)
5. Malgieri G, Comandé G. Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law*. 2017;7(4). (In Eng.)
6. Selbst AD, Powles J. Meaningful information and the right to explanation. *International Data Privacy Law*. 2017;7(4). (In Eng.)
7. Talia BG, Jann S. Big Data and Discrimination. *University of Chicago Law Abstract*. 2019;459. (In Eng.)
8. Wachter S, Mittelstadt B, Luciano 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;7(2). (In Eng.)