



**ОБЩЕРОССИЙСКАЯ ОБЩЕСТВЕННАЯ ОРГАНИЗАЦИЯ
«АССОЦИАЦИЯ ЮРИСТОВ РОССИИ»
МОСКОВСКОЕ ОТДЕЛЕНИЕ
Комиссия по правовому обеспечению Цифровой экономики**

119991, г. Москва, ГСП-1, Ленинские горы, д. 1, стр. 13,
4 учебный корпус МГУ имени М.В. Ломоносова, Юридический факультет, ауд. 611 а
тел.: +7 (499) 706-00-06, доб. 148; e-mail: digitallaw@alrf.msk.ru
<http://www.alrf.msk.ru>

«29» марта 2021 года

г. Москва

ЗАКЛЮЧЕНИЕ

**ПО ТЕМЕ «ПРАВОВЫЕ ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ПРОЗРАЧНОСТИ И
ЗАЩИТЫ ПРАВ ГРАЖДАН ПРИ ПРИНЯТИИ ЮРИДИЧЕСКИ-ЗНАЧИМЫХ
РЕШЕНИЙ В АВТОМАТИЗИРОВАННОМ (АЛГОРИТМИЧЕСКОМ) ПОРЯДКЕ В
УСЛОВИЯХ ЦИФРОВОЙ ЭКОНОМИКИ»**

Комиссия по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России (далее – Комиссия) провела ряд заседаний по теме «Правовые проблемы обеспечения прозрачности и защиты прав граждан при принятии юридически-значимых решений в автоматизированном (алгоритмическом) порядке в условиях цифровой экономики». По результатам заседаний Комиссия считает целесообразным опубликовать заключение с описанием ключевых проблем и предложений по вопросам, затронутым в докладах и рабочей дискуссии, и тем самым инициировать дальнейшее обсуждение данной проблематики с участием более широкого круга заинтересованных лиц.

ВВЕДЕНИЕ

Развитие технологий искусственного интеллекта и ускоренная пандемией коронавирусной инфекции цифровизация бизнес-процессов и сфер государственного управления поставили на повестку дня ряд новых проблем и вопросов. Одним из них является эффективная защита прав и законных интересов граждан и организаций при использовании государственными органами и организациями систем алгоритмического принятия решений, имеющих юридическое значение, без участия в процессе принятия такого решения человека

или при его номинальном участии (далее – системы АПР). Такие системы нередко имеют в своей основе технологии искусственного интеллекта, в том числе основанные на нейронных сетях и машинном обучении. Как следствие, они обладают немалым потенциалом по сокращению транзакционных издержек и повышению эффективности управленческих решений. Вместе с тем, такие системы несут в себе риски дискриминации граждан и организаций, вследствие:

- возможного наличия предубеждений (*bias*) в алгоритмах¹, умышленно или бессознательно привнесенных разработчиками; или
- некачественных или неполных данных, используемых для обучения системы или принятия решений в конкретной ситуации.

В банковской сфере система АПР может повлечь отказ гражданину в выдаче кредита по причине квалификации его алгоритмами системы в качестве неблагонадежного, поскольку у него в друзьях в социальных сетях фигурируют лица, не возвращавшие кредит ранее, либо потому, что лингвистический анализ публикуемых таким гражданином постов или фотографий демонстрирует по мнению разработчиков алгоритма, невысокий социальный статус такого лица. Никакого внешнего контроля над обоснованностью таких выводов алгоритмов нет, это исключительное усмотрение их разработчика. Проблема усугубляется и тем, что входные данные, используемые для работы системы могут быть неполными, неточными или «фейковыми». Таким образом, непрозрачность алгоритмов и источников данных для принятия решений, отсутствие реальной возможности у гражданина как-то влиять на процесс принятия на основе АПР решений в совокупности с их масштабным использованием в банковской сфере означает, что он не может получить ипотеку, потребительский кредит и реализовать многие свои насущные потребности.

В страховой сфере системы АПР могут использоваться для анализа страховых рисков в отношении данного конкретного страхователя – гражданина или организации. Его манера вождения, хобби и увлечения, медицинские данные и прочие факты будут взвешены алгоритмом с установлением индивидуализированной страховой премии, которую многие не смогут себе позволить заплатить. В итоге страхование может стать доступным только для тех случаев, где с точки зрения страховой компании риски минимальны, что деформирует саму идею страхования.

В сфере трудоустройства системы АПР, используемые рекрутерами, осуществляют первичную фильтрацию резюме, исходя из непрозрачных критериев, заложенных в

¹ Под алгоритмами понимается набор правил или инструкций, которым следуют при выполнении вычислений, или, в более общем смысле, набор процедур решения проблем, выполнение которых дает определенный результат. W. Barfield, J. Barfield. *An Introduction to Law and Algorithms. – The Law of Algorithms* (ed. by W. Barfield). – 2021. Cambridge University Press. – P. 4.

алгоритмах системы их разработчиком. В итоге многие резюме HR-сотрудники даже не видят, что не может не влиять на фактическую реализацию гражданами прав на трудоустройство.

Выявление дискриминационных начал в системах АПР крайне затруднено вследствие отсутствия правовых норм, регламентирующих возможность получения доступа независимых экспертов к алгоритмам и наборам данных, используемых для принятия решений, а также наличия правовых барьеров, связанных с законодательством о защите интеллектуальной собственности разработчиков таких систем. В свою очередь, обжалование решений, вынесенных такими системами, сопряжено с рядом сложностей, в частности:

- непрозрачностью самого алгоритма принятия решений (принцип «черного ящика»), вследствие чего сложно проанализировать причины принятия системой такого решения;
- отсутствия удобных для граждан и оперативных механизмов обжалования такого решения с достаточным количеством компетентных сотрудников оператора системы.

В настоящее время в Российской Федерации отсутствует законодательное регулирование использования систем АПР. Существует только ограниченный набор общих норм, которые потенциально могут быть применены к данным отношениям и то, при условии их «творческой интерпретации». В этой связи Комиссия обратила особое внимание на зарубежный опыт регулирования данных вопросов, поскольку отсутствие комплексного законодательного регулирования рассматриваемого вопроса дает России широкий простор по оценке и имплементации международного опыта правового регулирования (без механического копирования) в законодательство РФ с учетом российских реалий и особенностей правовой системы.

ЗАРУБЕЖНЫЙ ОПЫТ И ПРАКТИКА РЕГУЛИРОВАНИЯ

Специальное регулирование отношений, связанных с АПР, существует далеко не во всех странах. Многообразие алгоритмов и сфер применения систем АПР, новизна отношений и множество стейкхолдеров в данной области пока не позволили выработать детальные правила поведения в этой области.

В более чем 50 странах приняты рамочные документы и стратегии развития искусственного интеллекта (ИИ), которые затрагивают такие основные вопросы как:

- права человека и его роль в управлении алгоритмами;
- качество предоставляемых данных;
- прозрачность, предсказуемость и объяснимость алгоритмов;
- многообразие и справедливость;
- ответственность и подотчетность.

Второе заметное направление регулирования – это попытки устранить отдельные правовые и технологические барьеры, которые препятствуют более широкому применению искусственного интеллекта в принятии юридически-значимых решений.

При этом следует особо выделить следующие страны.

1. Канада

Канада была первой страной, принявшей в 2017 году национальную стратегию в области искусственного интеллекта – пятилетний план по инвестированию в исследования в области ИИ. Канадская стратегия ИИ отличается от других стратегий, потому что, в первую очередь, нацелена на проведение научных исследований. Для этого должны быть созданы новые институты по исследованию ИИ и отраслевые кафедры, что приведет к повышению международного авторитета Канады как лидера в этой области.

Также Канада стала одной из первых стран, которая приняла специальный закон, посвященный регламентации порядка использования АПР – Директиву о принятии решений в автоматизированном порядке², которая вступила в силу 1 апреля 2020 года. Указанная директива применяется к системам АПР (в том числе статистическим моделям), эксплуатируемым в сфере государственного управления. Данный законодательный акт является в настоящее время одним из наиболее детальных и проработанных, в связи с чем заслуживает достаточно подробного изложения в рамках данного заключения.

В основе Директивы лежит риск-ориентированный подход, который предусматривает классификацию систем АПР на 4 категории, в зависимости от степени возможного воздействия на права граждан и сообществ, их здоровье и благополучие, экономические интересы индивидов и организаций, устойчивое развитие экосистем.

В зависимости от класса системы АПР Директива предусматривает конкретные меры по минимизации рисков, связанных с нарушением прав и свобод граждан и организаций при ее функционировании.

1) Размещение уведомления о применении системы АПР на соответствующем сайте государственного органа.

Для систем, использование которых сопряжено с высокими рисками и долгосрочными негативными последствиями, такое уведомление должно содержать:

- описание того, как компоненты системы АПР функционируют,
- как они обеспечивают возможность принятия административного решения,
- результаты аудита или иных оценок такой системы, а также

² Directive on Automated Decision-Making. URL: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>

- описание массива обучающих данных или ссылку на сам такой массив в анонимизированном виде, если он опубликован в сети Интернет.

Системы АПР, использование которых сопряжено со средними рисками, требуют лишь уведомления об использовании такой системы в деятельности государственного органа при предоставлении соответствующей услуги.

2) Необходимость обеспечения независимой внешней оценки (peer review) системы определенным классом экспертов.

Данное требование предусматривается для большинства систем.

В качестве экспертов могут выступать либо специалисты из иных государственных органов, либо из неправительственных организаций или научных учреждений, либо привлеченные на договорной основе коммерческие организации, обладающей экспертизой в данной области.

В качестве альтернативы привлечению внешнего эксперта предусматривается возможность публикации спецификаций системы АПР в журнале с независимой системой рецензирования (*peer reviewed journal*).

Возможно также создание для оценки системы специального консультативного совета по правилам, указанным Министерством Финансов Канады.

Для наиболее опасных систем АПР необходимо сочетание как минимум двух из таких мер для обеспечения большей степени прозрачности и качества экспертной оценки.

3) Обеспечение участия человека в принятии решения.

Так, системы АПР, использование которых сопряжено с высокими рисками и долгосрочными негативными последствиями, требуют обязательного участия человека в процессе принятия решения, в том числе путем делегирования ему права принятия финального решения. Например, системы АПР могут быть использованы в судебном процессе, но они не могут выносить решение по делу, особенно, по уголовному.

4) Постоянное обучение сотрудников по вопросам функционирования системы АПР.

В зависимости от уровня риска и тяжести потенциальных последствий, такое обязательное обучение может иметь форму:

- изучения документации о такой системе, описывающей ее функциональность и дизайн;

- прохождения сотрудниками специальных обучающих курсов в дополнение к изучению документации;

- регулярного прохождения таких курсов с обеспечением возможности верификации факта их прохождения.

5) *Наличие запасного плана на случай сбоев и иных непредвиденных ситуаций при функционировании системы АПР.*

Такого рода план обязателен для систем, использование которых сопряжено с высокими рисками и долгосрочными негативными последствиями, чтобы граждане могли реализовать свои права даже в случае сбоев в функционировании системы. Как правило, такой план может быть реализован путем создания резервных систем.

6) *Определенный уровень административного согласования решения о введении в эксплуатацию системы.*

Данное требование направлено на обеспечение должного уровня ответственности должностных лиц при принятии решений об использовании систем, использование которых сопряжено с высокими рисками и долгосрочными негативными последствиями. В зависимости от класса системы, одобрение может быть получено у заместителя руководителя государственного органа, в котором система будет использоваться, а иногда только у Отдела Казначейства Тайного Совета Королевы (*Treasury Board of the Queen's Privy Council*) или должностных лиц еще более высокого уровня.

7) *Требование о предоставлении достаточных объяснений в отношении полученного в рамках системы решения.*

В зависимости от класса системы, такое объяснение может предоставляться:

- в общем виде в секции «Часто задаваемые вопросы» соответствующего раздела сайта государственного органа, эксплуатирующего систему;

- в форме предоставляемого по запросу объяснения в отношении принятого решения об отказе в предоставлении определенного блага, услуги или иного затребованного действия;

- в обязательном порядке, сопровождая любое отрицательное решение вне зависимости от наличия запроса от субъекта.

8) *Обеспечение адекватных механизмов обжалования.*

Директива не конкретизирует, как именно должны быть реализованы такие механизмы.

Одной из важнейших мер, направленных на обеспечение прозрачности при функционировании системы АПР, является обеспечение при определенных условиях доступа к исходному коду системы для проведения внешнего аудита.

По общему правилу (ст. 6.2.6 Директивы), если такого рода система является разработкой, сделанной по заказу государственного органа Канады, ее исходный код должен быть сделан общедоступным по лицензии типа *open source*, за исключением случаев, когда такой программный код обрабатывает данные, относящиеся к секретным, особо секретным или защищенным. Кроме того, исключение из требования публичности исходного кода может

быть в индивидуальном порядке предоставлено по решению Главного директора по информации Канады (*Chief Information Officer of Canada*).

Если в системе АПР используются проприетарные компоненты, принадлежащие третьим лицам, то такие компоненты должны быть предоставлены соответствующему государственному органу – оператору системы.

Иными словами, облачное предоставление доступа к системе АПР, используемой в сфере государственного управления, при котором провайдером выступает частное лицо, без одновременного физического предоставления экземпляров программных компонентов государственному органу, является недопустимым.

Правительство Канады должно иметь доступ к ним в объеме, необходимом для проведения аудита, расследований, судебных и административных процедур, при соблюдении гарантий защиты таких компонентов от несанкционированного доступа. Кроме того, Правительство Канады сохраняет за собой право предоставить полномочия третьим лицам по осуществлению экспертизы и аудита таких проприетарных компонентов.

2. Европейский Союз

На уровне общеевропейского законодательства в настоящее время отсутствуют специализированные директивы или регламенты, посвященные проблематике использования систем АПР. Если использование систем АПР предполагает обработку персональных данных, минимальные юридически-обязательные требования закреплены в Общем регламенте защиты персональных данных (*General Data Protection Regulation, GDPR*):

1) По общему правилу, использование систем АПР, влекущее за собой юридически-значимые последствия, допустимо только в случаях, когда это прямо допускается применимым законодательством страны-оператора, необходимо для заключения или исполнения договора с субъектом персональных данных либо с согласия субъекта.

2) У субъекта персональных данных должна быть возможность получить информацию об общей логике устройства применяемой системы АПР и потребовать вмешательства человека, чтобы он мог изложить свои возражения и оспорить вынесенное решение.³

При этом отсутствие детального и универсального общеевропейского законодательства дает широкие возможности отдельным странам членам ЕС регулировать системы АПР на уровне судебной практики, применяющей общие положения национального законодательства.

³ Статьи 21 и 22 и пункты (71) и (72) преамбулы GDPR

Например, во **Франции** Кодекс взаимоотношений между правительством и обществом (Code des relations entre le public et l'administration)⁴ предусматривает, что любое решение, вынесенное с использованием алгоритмов, должно содержать указание на данный факт. Помимо этого, законодательство обязывает по требованию предоставлять информацию о том, на каких правилах основан алгоритм и в чем заключается его роль.⁵ Более того, в июне 2018 г. Конституционный совет Франции (Conseil constitutionnel) указал, что решение, основанное исключительно на системе АПР, правомерно только в том случае, если алгоритм и его устройство могут быть полностью объяснены лицу, в отношении которого вынесено решение.⁶ Прозрачность алгоритмов рассматривается как форма контроля. Закон «О цифровой республике» (Loi pour une République numérique)⁷ 2016 г. указывает, что правила, определяющие основные алгоритмические обработки, должны публиковаться на сайте государственного органа.⁸

В **Нидерландах** суды активно изучают границы допустимого использования систем АПР в государственном управлении и судопроизводстве. Например, в феврале 2020 г. окружной суд Гааги Rechtbank Den Haag вынес решение о том, что алгоритмический инструмент, используемый правительством для обнаружения мошенничества в налоговой сфере и сфере социального обеспечения, нарушает требования Европейской конвенции по правам человека.⁹

В **Германии** можно отметить дело Верховный суд Германии отклонил иск клиента банка о предоставлении информации об алгоритмах банка, который принимает решение об одобрении кредита, к кредитной организации, поскольку счел это охраняемой коммерческой тайной. Однако Суд разрешил клиентам банков запрашивать информацию относительно данных, используемых для расчета кредитоспособности с помощью алгоритмов¹⁰.

3. Соединенные Штаты Америки

Американское законодательство не содержит специальных норм, посвященных регулированию использования систем АПР. Причиной особого подхода США к вопросам регулирования персональных данных, которое в определенной степени ретранслируется и на проблематику регулирования ИИ и систем АПР, является вера в возможности «мягкого права»

⁴ https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000031366350/2021-02-24/

⁵ https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205535/

⁶ <https://www.nextinpact.com/article/28508/106743-transparence-algorithmes-publics-lavertissement-conseil-constitutionnel>

⁷ <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000033202746/>

⁸ 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.*

⁹ <https://www.loc.gov/law/foreign-news/article/netherlands-court-prohibits-governments-use-of-ai-software-to-detect-welfare-fraud/>

¹⁰ <https://openjur.de/u/677956.html>

и саморегулирования, а также в то, что рынок «все расставит по местам». Не последнюю роль играет и позиция американских корпораций, активно лоббирующих свои интересы в американском правительстве.

В качестве примера документа, представляющего собой «мягкое право», можно привести проект Руководства по регулированию приложений ИИ ¹¹, подготовленный в 2020 году Управлением по политике в области науки и технологий Белого дома. Документ содержит свод принципов, которые необходимо учитывать государственным органам при формировании регулирования использования технологий ИИ в частном секторе. Данный документ релевантен и для систем АПР, так как включает в себя десять принципов, которые регуляторам необходимо учитывать при принятии решения о том, следует ли регулировать ИИ и, если да, то каким образом:

1. Доверие общественности к ИИ.

2. Участие общественности.

3. Научная целостность и качество информации.

4. Оценка и управление рисками. Отмечено отсутствие необходимости исключать все предсказуемые риски. Вместо этого следует использовать подход, основанный на оценке риска, чтобы определить, какие риски оправданы, а какие потенциально могут нанести непоправимый вред или ущерб (т.е. затраты на ликвидацию последствий превышают ожидаемые выгоды от использования ИИ).

5. Выгоды и затраты. Регуляторы должны тщательно оценить затраты, выгоды и воздействие использования ИИ перед разработкой нормативных требований, особенно что касается уточнения законодательства в части ответственности за решения, принимаемые ИИ;

6. Гибкость правил, поскольку жесткие правила в большинстве случаев будут непрактичными и неэффективными.

7. Справедливость и недискриминация в отношении результатов и решений, принимаемых приложениями ИИ.

8. Раскрытие информации и прозрачность, что может способствовать повышению общественного доверия.

9. Безопасность.

10. Сотрудничество. Регуляторы должны координировать свои действия друг с другом, чтобы делиться опытом и обеспечивать согласованность политики и действия в области использования ИИ.

¹¹ <https://www.insidetechnia.com/2020/01/14/ai-update-white-house-issues-10-principles-for-artificial-intelligence-regulation/>

Вместе с тем среди регуляторов и законодателей США появляется понимание о невозможности решения всех проблем, связанных с широкомасштабным использованием систем АПР в различных сферах общественной жизни исключительно на основании саморегулирования и рекомендаций. В этой связи в Конгрессе США и законодательных органах отдельных штатов появляются первые законопроекты в указанной сфере, в которых отражены различные инструменты, направленные на обеспечение прозрачности при использовании АПР.

Одним из первых законопроектов, внесенных в Конгресс США и посвященных проблематике АПР, является проект Закона о подотчетности алгоритмов (*Algorithmic Accountability Act of 2019*)¹². Данный законопроект имеет своей целью наделение Федеральной торговой комиссии США полномочиями по регулированию порядка использования определенными лицами алгоритмов, которые имеют в своей основе обработку персональных данных граждан и могут негативно влиять на их права и свободы. В качестве обязанных лиц фигурируют, в частности, организации, выручка которых за 3 года составляет более \$ 50 млн., либо которые обрабатывают персональные данные более 100 000 потребителей или данные с более, чем 1 млн. потребительских устройств. В качестве основных инструментов обеспечения прозрачности использования данных алгоритмов законопроект устанавливает обязанности осуществления Оценки алгоритмического воздействия (*Algorithmic Impact Assessment*), предполагающую оценку системы АПР на предмет справедливости, отсутствия дискриминации, точности и релевантности обучающих данных, а также Оценки воздействия на защиту персональных данных (*Data Protection Impact Assessment*) на предмет степени защищенности данных от утечек и иных уязвимостей, а также степени реализации права на неприкосновенность частной жизни. Законопроект был внесен еще в 2019 г. и все еще находится на ранних стадиях рассмотрения. В случае принятия его действие может затронуть не только технологические компании, разрабатывающие алгоритмы, но и банки, страховые организации, розничный бизнес и прочие компании, которые используют в своей деятельности системы АПР. Такие организации должны будут проактивно предпринимать меры и должным образом документировать их по минимизации рисков от их использования под надзором регулятора.

Отдельные штаты США, не дожидаясь появления федерального законодательства, сами занимаются активным регулированием в сфере вопросов, касающихся порядка использования систем АПР.

¹² <https://www.congress.gov/bill/116th-congress/house-bill/2231/all-actions?overview=closed#tabs>

Наибольшую активность в этом вопросе проявляет штат Нью-Йорк. Данный штат стал первым в США, где специальным законом была создана особая рабочая группа, на которую возложены задачи по обеспечению прозрачности и справедливости при применении систем АПР в деятельности органов власти штата – *ADS Task Force*¹³. Данная рабочая группа включает в себя не только представителей органов власти, но и представителей научного сообщества, правозащитных организаций¹⁴. В число задач, поставленных перед данным органом, входит анализ используемых органами власти штата Нью-Йорк систем АПР и подготовка рекомендаций мэру о необходимости имплементации дополнительных процедур, направленных на минимизацию рисков дискриминации, повышения уровня обратной связи системы, обеспечения информированности граждан об использовании таких систем, целесообразности обеспечения архивирования принятых системами АПР решений и входных данных, использованных для их принятия.

В тех штатах, в которых отсутствует специальное законодательство, границы допустимого использования алгоритмических инструментов в принятии юридически-значимых решений определяет судебная практика. Например, в 2016 г. в деле *State v. Loomis* Верховный суд штата Висконсин признал допустимым использование алгоритмических инструментов COMPAS при осуществлении правосудия, а именно при оценке вероятности рецидива. При проведении предварительного расследования Исправительный департамент штата Висконсин обращался к системе COMPAS, оценивающей риск и прогнозирующей рецидивизм. Программа проанализировала всю информацию, и Лумис был приговорен к 6 годам лишения свободы, после чего он подал иск в Верховный суд штата Висконсин¹⁵ с требованием пересмотреть использование алгоритма для вынесения судебных решений. По мнению Лумиса, система работает некорректно и выносит решения предвзято, основываясь на общих фактах (пол и возраст). Кроме того, алгоритм применяет секретные критерии оценки личности подсудимого, суть которых не разглашается, поскольку данный алгоритм представляет собой коммерческую тайну.

Однако в 2016 году Верховный суд посчитал, что надлежащая правовая процедура не была нарушена, поскольку решение было вынесено судьей, а алгоритм лишь подтвердил намерение судьи вынести меру пресечения в форме лишения свободы сроком на 6 лет. Суд отклонил три основных довода защиты. Во-первых, защита указывала на нарушение принципа «справедливого судебного разбирательства» (*due process*), т.к. защите не было предоставлено возможности проверить ни особенности работы алгоритма, ни актуальность и точность

¹³<https://www1.nyc.gov/office-of-the-mayor/news/251-18/mayor-de-blasio-first-in-nation-task-force-examine-automated-decision-systems-used-by>

¹⁴ NYC Automated Decision Systems Task Force at <https://www1.nyc.gov/site/adstaskforce/index.page>.

¹⁵ <https://caselaw.findlaw.com/wi-supreme-court/1742124.html>

сведений, на которые он опирался. Во-вторых, защита настаивала на том, что использование инструмента COMPAS нарушает право подсудимого на индивидуальное решение, т.к. инструмент предназначен для работы с коллективными, а не личными данными. В-третьих, защита указывала на дискриминацию по признаку пола в связи с тем, что расчетные показатели учитывали пол обвиняемого. Верховный суд штата отклонил каждый из трех доводов защиты, ограничившись символическим требованием, чтобы нижестоящие суды указывали факторы, принятые ими во внимание, помимо результатов оценки COMPAS.¹⁶ Лумис не смог доказать, что оценка алгоритма была решающим фактором в избрании меры пресечения.

В 2017 году Лумис обратился в Верховный суд США с приказом об истребовании дела (writ of certiorari)¹⁷, но ему было отказано в пересмотре дела.

Система COMPAS используется более чем в 10 штатах и с ее помощью судьи определяют меру пресечения. Еще до указанного прецедента алгоритм подвергался критике в связи с его предвзятостью и необъективностью. Союз гражданских свобод США отмечал предвзятость алгоритма и учет ошибочных данных, предоставленных органами правопорядка. Издание ProPublica опубликовало исследование¹⁸ использования алгоритма COMPAS при рассмотрении 10 тыс. дел в штате Флорида, подтверждая ошибочное присуждение высокого риска рецидивизма чернокожим преступникам.

В деле *State of Kansas v. John Keith Walls* суд указал, что содержание права на объяснение работы алгоритма должно интерпретироваться более широко, то есть с акцентом не на само решение, а на объяснение условий, в которых оно принималось, и его исходных данных. Получение данного представления становится возможным при предоставлении операторами систем доступа к ней и ее исходному коду и при объяснении основ работы алгоритма. Госорганы не заинтересованы в такой открытости. По мнению государственных структур раскрытие внутренней информации о работе алгоритмов несет в себе угрозы для их безопасной работы. Суд принял решение в пользу гражданина.¹⁹

4. Великобритания

В Великобритании отсутствуют законы, специально посвященные регулированию вопросов использования систем АПР. Однако многие аспекты такого использования

¹⁶ 881 N.W.2d 749 (Wis. 2016), Han-Wei Liu, Ching-Fu Lin, Yu-Jie Chen, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, *International Journal of Law and Information Technology*, Volume 27, Issue 2, Summer 2019, Pages 122–141, <https://doi.org/10.1093/ijlit/eaz001>

¹⁷ <https://www.supremecourt.gov/docketfiles/16-6387.htm>

¹⁸ <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

¹⁹ *The Court of Appeals of the State of Kansas. State of Kansas v. John Keith Walls. Case no. 116,027. 7 November 2017* <https://pittsreporting.kscourts.org/Appellate/CaseDetails?caseNumber=116027>

подпадают под действующие нормы законодательства о недопустимости дискриминации и защите персональных данных. Важное значение имеют также административно-правовые нормы, выраженные в прецедентах об обжаловании действий и решений, принятых государственными органами, и разъяснения регуляторов, посвященные регулированию применения технологий искусственного интеллекта и систем АПР с учетом формирующихся «лучших практик».

Так, например, Управление комиссара по информации (Information Commissioner's Office, ICO) совместно с Институтом Алана Тьюринга ²⁰ выпустили разъяснения по обеспечению прозрачности систем АПР и объяснению решений, принятых с использованием ИИ (далее – Руководство)²¹. Данный документ был подготовлен во исполнение поручения английского правительства, сформулированного в рамках программы развития ИИ в Великобритании (AI Sector Deal, 2018). Он логически развивает положения и рекомендации, изложенные годом ранее в Руководстве ICO по применению положений GDPR по обеспечению прав граждан при принятии касающихся их решений исключительно на основании автоматизированной обработки данных, в том числе посредством профайлинга.²²

Данный документ состоит из трех частей.

Часть 1 Руководства посвящена Основам объяснения ИИ, адресованным Ответственным за обработку персональных данных в организациях (Data Protection Officers, DPO) и иным лицам, обеспечивающим соблюдение данного законодательства в организациях. При этом отмечается опасность предоставления чрезмерной информации в рамках выполнения организацией обязанности по объяснению, например, в виде сведений об алгоритмах, раскрывающих секреты производства, что может привести к жульническим и мошенническим действиям со стороны пользователей и попыткам «обыграть» алгоритм. Вместе с тем отмечается, что предоставление объяснений является не только выполнением требований законодательства, но и инструментом обеспечения доверия пользователей к системам ИИ и организации в целом за счет обеспечения прозрачности их функционирования, реальной возможности обжалования принятых решений, а также получения «обратной связи» в виде изменения их поведения в лучшую сторону.

²⁰ Управление комиссара по информации (Information Commissioner's Office, ICO) является уполномоченным органом по защите прав субъектов персональных данных в Великобритании. Одной из функций ICO является разработка руководств и рекомендаций по вопросам, входящим в компетенцию Управления.

²¹ *Explaining decisions made with AI, ICO Guidance, May 2020.* <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-ai/>

²² *Rights related to automated decision-making including profiling, ICO Guidance, May 2019.* <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/>

В этой связи выделяется 6 основных типов объяснений, которые организация может использовать для предоставления субъектам:

1) Логическое обоснование принятого решения, с указанием основной причины его принятия.

2) Кто несет ответственность за разработку, управление и имплементацию системы с использованием ИИ.

3) Сведения о данных, использованных для принятия решения, и как именно они были использованы.

4) Сведения о мерах, реализованных на этапе проектирования и разработки системы для обеспечения справедливости принимаемых решений.

5) Сведения о мерах, реализованных на этапе проектирования и разработки системы для обеспечения ее надежности и безопасности.

6) Сведения о мерах, реализованных на этапе проектирования и разработки системы для обеспечения мониторинга последствий принятия решений.

Руководство в деталях описывает какие именно сведения могут быть предоставлены в рамках каждого типа объяснений, а также когда такие объяснения должны предоставляться в виде предварительных уведомлений или сопровождать уже принятое решение.

Часть 2 Руководства адресована техническим специалистам и разъясняет для каких ситуаций какой тип объяснений подходит наилучшим образом, а также содержит пошаговые инструкции по реализации соответствующего типа пояснений. При этом особо отмечается, что проработка данных вопросов должна быть инициирована на ранних стадиях проектирования или закупки систем АПР. В приложениях к данной части Руководства приводятся примеры реализации рекомендаций в отношении конкретных систем АПР.

Часть 3 Руководства адресована менеджменту организации и акцентирует внимание на том, как следует адаптировать внутренние процедуры и политики компании, как распределить роли сотрудников при использовании систем АПР и организовать их обучение.

В июне 2020 года ISO разработало Руководство по госзакупкам ИИ²³ – инструкции, как покупать технологии ИИ, а также советы по решению проблем, которые могут возникнуть в процессе госзакупок. Документ рекомендует ответственно подходить к использованию алгоритмов и обеспечить их прозрачность путем объяснения принципов работы системы АПР. В этой связи необходимо отказаться от использования эффекта «черного ящика» и при проведении тендера требовать раскрытия принципов и механизмов работы ИИ в объеме,

²³https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/945884/Guidelines_for_AI_procurement.pdf

необходимом для понимания процесса принятия решений алгоритмом, основанным на использовании технологий ИИ.

Кроме того, ИСО ведет работу по разработке системы аудита ИИ.²⁴ Выделяют два ключевых компонента рамочных положений, касающихся аудита ИИ:

- 1) управление и подотчетность;
- 2) определение зон риска, связанных с использованием ИИ (понимание возможных рисков, связанных с защитой персональных данных, которые могут возникнуть в областях, специфичных для ИИ, и указание адекватных методов управления рисками).

Определено восемь областей риска, связанных с ИИ, среди которых, например:

- 1) справедливость и прозрачность профилирования (включая вопросы объяснимости самих алгоритмов ИИ и объяснимости принимаемых решений);
- 2) точность (в первую очередь имеется в виду точность используемых данных);
- 3) полностью автоматизированные модели принятия решений (проверка таких моделей и степень вмешательства человека); и др.

Из недавней судебной практики стоит отметить поданный 16 декабря 2020 года британскими водителями иск к Uber. Говорится, что Uber распределяет поездки и их тарифы на основе информации о характеристиках, поведении и других личных качествах водителей, которые собираются через приложение. Истцы требуют полный доступ к этой информации и работе алгоритмов, ссылаясь на то, что «Uber должен обеспечивать полную прозрачность».²⁵ На момент подготовки данного заключения спор еще не был завершен.

РОССИЙСКАЯ ФЕДЕРАЦИЯ

В России пока достаточно мало информации об использовании таких систем конкретными организациями или органами власти. Отчасти это связано с тем, что они сами не осознают, что используют такого рода системы, поскольку это воспринимается ими как одна из многих информационных систем или программ для ЭВМ. С другой стороны, они не хотят афишировать использование таких систем, чтобы не провоцировать лишних вопросов со стороны регуляторов и граждан. Но даже несмотря на это уже можно найти примеры применения систем АПР в России. Так, например, банки ВТБ²⁶ и Московский кредитный банк²⁷ используют системы АПР при принятии решений по кредитам. В сфере

²⁴<https://ico.org.uk/about-the-ico/news-and-events/ai-blog-an-overview-of-the-auditing-framework-for-artificial-intelligence-and-its-core-components/>

²⁵ <https://www.reuters.com/article/netherlands-uber-lawsuit/british-drivers-take-uber-to-dutch-court-over-data-use-idINL1N2IW0SA>

²⁶ <https://lprime.ru/banks/20210128/832911901.html>, <https://mkb.ru/news/25561>

государственного управления системы АПР используются для привлечения к административной ответственности на основе данных видео- или фотофиксации, а также на основе иных технических данных, например, от сотовых операторов. Такого рода системы применялись для привлечения к ответственности весной 2020 г. в г. Москва за нарушение режимов «самоизоляции» и «карантина». В СМИ фигурировало много курьезных случаев ошибочного привлечения к ответственности, например, парализованных граждан или граждан, чья квартира находилась на границе Москвы и Московской области. Очевидно, что в данном случае речь идет и о некачественных данных, и о некорректных алгоритмах работы таких систем. МВД планирует реализовать пилотные проекты по выявлению признаков серийных (взаимосвязанных) преступлений и по определению индивидуальных анатомических признаков человека, полученных из биологического материала с мест совершения преступления. Запланировано формирование дата-сетов преступлений и правонарушений, а также иных обращений граждан и дата-сетов биологического материала с мест совершения преступления.²⁸

В любом случае, даже если предположить, что их использование в России в настоящее время не столь велико, как за рубежом, это никак не снимает остроты вопроса о целесообразности их адекватного правового регулирования. Поскольку их повсеместное применение в России – это лишь вопрос времени, так как это вопрос технологии и экономики, которые универсальны для большинства стран, безотносительно их культурных и страновых особенностей.

ВЫВОДЫ И РЕКОМЕНДАЦИИ

1. Представляется ошибочным чисто технократический подход к внедрению систем ИИ и АПР, направленный на максимальное снятие барьеров к их внедрению. Такой подход создает условия для их повсеместного распространения и бесконтрольного использования, при котором эффективность запоздавшего регулирования будет крайне низкой. В данном случае грамотная превенция является предпочтительнее, чем устранение масштабных негативных последствий их применения. Как свидетельствует зарубежный опыт, некоторые сферы использования таких систем могут и должны быть ограничены в случае их высокой рискованности.

Примером сферы, в которой целесообразно рассмотреть направления возможных ограничений по разработке и применению систем поддержки принятия решений является сфера судопроизводства. Так, например, Франция стала первой страной в мире, которая прямо

²⁸ https://www.cnews.ru/articles/2021-02-15_v_rossii_budut_raskryvat_prestupleniya

запретила под страхом уголовной ответственности исследования индивидуального поведения судей. Уголовно-наказуемым деянием является «оценка, анализ, сравнение или прогнозирование» поведения отдельных судей. Это обусловлено тем, что подобного рода аналитика способствует так называемому «форум-шоппингу» со стороны участников процесса, а также способна влиять на беспристрастность судей. Иными словами, такого рода системы могут оказывать деформирующее воздействие на судебную систему. Данные опасения вполне созвучны и в отношении российской судебной системы, «алгоритмизация» которой может негативным образом оказать влияние на принцип равенства сторон судопроизводства и предоставить необоснованные преимущества отдельным компаниям, разрабатывающим и применяющим соответствующие технологические решения.

По мнению Комиссии АПР в судопроизводстве может вводиться только после его апробирования и в «цифровых песочницах» в рамках экспериментально-правовых режимов и только в отношении определенных категорий споров, которые не связаны с установлением множества фактов и необходимостью оценки разнородных доказательств. Как вариант, в отношении тех споров, которые рассматриваются в порядке приказного производства или в порядке упрощенного производства.

2. Несколько преждевременным представляется введение универсального регулирования использования систем АПР, распространяемого и на коммерческие организации, и на публичный сектор.

Многие зарубежные правовые порядки идут по пути регламентации использования таких систем именно в публичном секторе. Данный подход видится обоснованным, так как именно в сфере государственного управления могут приниматься наиболее чувствительные решения в отношении граждан с одной стороны, а с другой стороны, государство является вполне логичным кандидатом на апробирование результатов выработанного им регулирования.

По итогам проверки в государственных структурах подходы, доказавшие свою эффективность с точки зрения защиты прав и свобод граждан, могут быть масштабированы на коммерческий сектор. При этом важно обеспечить прозрачность систем АПР, их объяснимость и аудируемость, а также соблюдение требований к качеству входящих данных. В частности, целесообразно, по примеру Канады, рассмотреть вопрос о введении механизмов независимой экспертной оценки на этапе закупки таких систем и их имплементации, а также повышения уровня принятия решений об имплементации таких систем (как вариант, на уровень руководства министерства или аналогичного по уровню ведомства - для противодействия самодеятельности чиновников на местах, минимизации коррупциогенных факторов и повышения степени подотчетности государственных органов за результаты использования соответствующих систем. Крайне важным также является создание

эффективной системы обжалования принятых в автоматизированном порядке решений государственных органов. Такая система должна быть основана на качественно иных принципах, нежели механизмы обжалования традиционных «аналоговых» решений. Она должна быть максимально «дружелюбной» для граждан и организаций, оперативной и обеспечивающей участие компетентных лиц, рассматривающих жалобу. Следует предусмотреть механизмы, направленные на минимизацию стимулов по автоматическому «штампованию» вынесенных системами решений указанными лицами, а также обеспечивающих совершенствование алгоритмов системы по результатам рассмотрения жалоб («обеспечение обратной связи»).

3. Использование систем АПР в коммерческом секторе нуждается в адекватной регламентации. Немалую роль в этой части должны играть механизмы саморегулирования. Уже сейчас существуют примеры достаточно успешных инициатив в этой области. Однако вряд ли можно рассчитывать, что этические кодексы и саморегулирование смогут сами по себе справиться с защитой граждан и организаций от возможных негативных аспектов, поскольку главной целью любой коммерческой организации является извлечение прибыли, что накладывает отпечаток на мотивы и характер ее действий. В этой связи правовое регулирование должно подкреплять положения этических кодексов и действовать в тех случаях, когда саморегулирование не срабатывает эффективным образом. Представляется целесообразным максимальное использование уже существующих норм в законодательстве с учетом их разъяснений компетентными органами. В качестве примера можно привести положения ч. 1 ст. 16 Федерального закона РФ от 27 июля 2006 года № 152-ФЗ «О персональных данных» (далее - 152-ФЗ), в которой установлен общий запрет на «принятие на основании исключительно автоматизированной обработки персональных данных решений, порождающих юридические последствия в отношении субъекта персональных данных или иным образом затрагивающих его права и законные интересы». Исключениями из данного запрета являются два случая, указанные в ч.2 ст. 16 152-ФЗ:

- 1) наличие согласия в письменной форме субъекта персональных данных;
- 2) указание федерального закона, устанавливающего также меры по обеспечению соблюдения прав и законных интересов субъекта персональных данных.

Данная норма, несмотря на почти пятнадцатилетний срок существования, не породила какой-либо значимой правоприменительной практики. Вместе с тем в условиях существенного возрастания сферы применения систем алгоритмического принятия решений (выдача кредитов, принятие решений о приеме на работу или в учебное заведение), есть все основания для того, чтобы эти положения приобрели вторую жизнь. Они обладают достаточной гибкостью для того, чтобы стать фундаментом для обеспечения должного уровня

прозрачности при использовании таких систем во взаимоотношениях с гражданами, поскольку большинство ситуаций применения АПР, создающих риски для прав и свобод граждан, сопряжено именно с обработкой персональных данных. При этом целесообразным является совершенствование данных положений с учетом нововведений, привнесенных GDPR, а также правоприменительной практики европейских регуляторов. В частности, крайне важно уйти от формального толкования фразы о том, что решение принимается на основании «исключительно» автоматизированной обработки. Следует уточнить, как это сделано европейскими регуляторами, что формальное присутствие человека на каком-то этапе принятия решения не должно исключать применения соответствующих защитных положений закона, если такой человек при этом не наделен полномочиями влиять на содержание решения, принятого системой АПР и не может его изменить.²⁹

Данное уточнение позволит минимизировать риски обхода закона операторами систем АПР за счет включения человека для номинального участия на каком-либо этапе решения системой АПР, например, для оформления принятого системой решения в «бумажный вид». Исключения, предусмотренные ч. 2 ст. 16 152-ФЗ о персональных данных могут стать основой для разработки специального законодательства, регламентирующего использование систем алгоритмического принятия решений в отдельных сферах. Согласие субъекта, отвечающее требованиям 152-ФЗ (в том числе требования ч. 1 ст. 9 о том, что оно дается субъектом «свободно, своей волей и в своем интересе»), может стать основным инструментом легитимации использования систем алгоритмического принятия решений коммерческими операторами, не исключая однако возможность применения в отдельных случаях второго типа исключения в виде указания закона.

Второй тип легитимирующих основания для применения систем алгоритмического принятия решений все же по своей сути больше подходит для сферы государственного управления. При этом следует подчеркнуть, что в положении ч. 2 ст. 16 152-ФЗ о данном исключении содержится два принципиально важных момента:

1) установление возможности применения систем алгоритмического принятия решений с порождением юридически-значимых результатов только на основании федерального закона (таким образом, региональное и подзаконное правотворчество в этой области должно быть исключено);

2) в таком федеральном законе помимо общей легитимации применения систем алгоритмического принятия решений в конкретной сфере должны предусматриваться также и

²⁹ *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Working Party Article 29, October 2017. P. 21.*

конкретные «меры по обеспечению соблюдения прав и законных интересов субъекта персональных данных».

К числу таких мер могут относиться с учетом анализа зарубежного опыта:

- размещение уведомления о применении системы алгоритмического принятия решений, заложенных в нее принципах и наборах использованных данных на соответствующем сайте государственного органа или иной организации;

- определенный уровень административного согласования решения о введении в эксплуатацию системы алгоритмического принятия решений;

- осуществление предварительной оценки алгоритмического воздействия системы на предмет справедливости, отсутствия дискриминации, точности и релевантности обучающих данных (Algorithmic Impact Assessment);

- необходимость обеспечения независимой внешней оценки (peer review) системы определенным классом экспертов;

- наличие запасного плана на случай сбоев и иных непредвиденных ситуаций при функционировании системы алгоритмического принятия решений;

- формирование регулярных отчетов об использовании системы, а также возможности оперативно вмешаться и модифицировать ее.

- возможность обжалования, основанная на информации о принципах работы системы АПР и использованных наборах данных.

Как отмечалось выше, имплементация органами государственной власти данных мер, с учетом специфики сферы применения системы АПР и потенциальных рисков для граждан, может стать отправной точкой для выработки комплексного регулирования порядка использования таких систем на различных этапах их жизненного цикла и в различных сферах деятельности.

Эксперты, участвующие в подготовке настоящего заключения:

1. Журавлев Александр Викторович – председатель Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России, управляющий партнер юридической компании «ЭБР»;

2. Савельев Александр Иванович – заместитель председателя Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России, доцент факультета права НИУ ВШЭ;

3. Едидин Борис Александрович – заместитель председателя Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России,

к.ю.н., заместитель генерального директора по правовым вопросам Института развития интернета (ИРИ);

4. Перевалов Вадим Андреевич – член Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России;

5. Казанцев Ефим Александрович – ответственный секретарь и член Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России, к.ю.н., эксперт Moscow Digital School;

6. Орехович Александра Владимировна – член Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России, директор по правовым инициативам Фонда развития интернет-инициатив (ФРИИ);

7. Левашенко Антонина Давидовна – член Комиссии по правовому обеспечению цифровой экономики Московского отделения Ассоциации юристов России, руководитель Центра компетенций и анализа стандартов ОЭСР РАНХиГС, старший научный сотрудник ВАВТ Минэкономразвития России и РАНХиГС, эксперт Центра стратегических разработок.

Председатель Комиссии

А.В. Журавлев