Научная статья
УДК 34:004:346.1:006.44:004.8
EDN: https://elibrary.ru/oppobg
DOI: https://doi.org/10.21202/jdtl.2023.42

Достижение алгоритмической прозрачности и управление рисками информационной безопасности при принятии решений без вмешательства человека: правовые подходы

Анна Константиновна Жарова

Институт государства и права Российской академии наук
г. Москва, Российская Федерация

Ключевые слова

GDPR,
алгоритмическая прозрачность,
защита данных,
информационная безопасность,
информационные технологии,
искусственный интеллект,
конфиденциальность,
персональные данные,
право,
цифровые технологии

Аннотация

Цель: сравнение современных подходов в праве к использованию в процессе принятия решений программных кодов и алгоритмов, отвечающих принципам прозрачности и открытости, а также возрастающим требованиям к обеспечению безопасности персональных и иных больших данных, полученных и обработанных алгоритмическим путем.

Методы: основными методами исследования принципа прозрачности алгоритмизированного принятия решений являлись формально-юридический и сравнительный анализ правовых актов и международных стандартов информационной безопасности, содержащихся в них принципов и правовых конструкций.

Результаты: определено, что развитие области стандартизации информационной безопасности, включение в правовые акты требований о разработке информационных технологий, соответствующих принципам прозрачности и открытости применяемых алгоритмов, позволит минимизировать риски, связанные с неправомерными обработкой больших пользовательских данных и получением информации об их частной жизни; выявлены связанные с реализацией алгоритмической прозрачности предложения в области правового регулирования обработки данных; сформулированы рекомендации, с опорой на которые законодатель может решать задачу обеспечения открытости логики работы алгоритмов информационных технологий с учетом современных стандартов информационной безопасности.

Научная новизна: состоит в обосновании новых тенденций и формируемых в соответствии с ними правовых подходов, позволяющих раскрыть логику обработки данных цифровыми и информационными технологиями, на основе характеристики общеевропейских стандартов концепции конфиденциальности при проектировании новых цифровых и информационных технологий принятия решений и защиты данных, новых правовых требований, предъявляемых к системам искусственного интеллекта, включая требование об обеспечении алгоритмической прозрачности, критериев обработки персональных данных, а также больших пользовательских данных. При этом защита данных рассматривается как система правовых, технических и организационных принципов, направленная на обеспечение конфиденциальности персональных данных.

Практическая значимость: обусловлена необходимостью изучения передового отечественного и международного опыта защиты частной жизни пользователей цифровых и информационных технологий, а также законодательного обеспечения требований об использовании алгоритмов, отвечающих принципам прозрачности и открытости обработки персональных данных с учетом необходимости обеспечения конфиденциальности на всех этапах жизненного цикла их обработки, что позволит обеспечить непрерывность управления безопасностью.

Для цитирования

Жарова, А. К. (2023). Достижение алгоритмической прозрачности и управление рисками информационной безопасности при принятии решений без вмешательства человека: правовые подходы. *Journal of Digital Technologies and Law*, *1*(4), 973–993. https://doi.org/10.21202/jdtl.2023.42

Содержание

Введение

Все сложнее становятся модели информационных технологий, и все больше данных, связанных с человеком, они обрабатывают. Так, технологии интернета вещей собирают большой объем данных, в котором могут содержаться различные данные, включая большие пользовательские данные. У пользователя информационными технологиями (далее – ИТ) вызывает обеспокоенность невозможность контроля над действиями, осуществляемыми информационными технологиями, а также отсутствие понимания логики работы алгоритмов, обрабатывающих его данные, какие именно данные обрабатываются и каков конечный результат анализа данных. Желание понять критерии анализа обрабатываемых данных, а также необходимость обеспечения контроля за их перечнем привели законодателя к мысли о необходимости включения в правовые акты требования об использовании алгоритмов, отвечающих принципам прозрачности и открытости. Иными словами, необходимо раскрыть логику обработки данных информационными технологиями.

В такой научной области, как информатика, для описания прозрачности процессов, происходящих при функционировании информационных технологий, применяется термин «алгоритмическая прозрачность». В связи с необходимостью обеспечения защиты пользовательских данных этот термин был заимствован юридической наукой.

Термин «алгоритмическая прозрачность» в законодательстве Российской Федерации используется для описания регулирования отношений при применении систем искусственного интеллекта (далее – ИИ)[1]. Хотя ученые считают, что термин «алгоритмическая прозрачность» может применяться для описания более широкого круга отношений, выходящих за пределы функционирования ИИ (Кутейников и др., 2020; Гулемин, 2022).

В связи с такой многозначностью в первую очередь необходимо изучить понятие «алгоритмическая прозрачность», а далее на основании полученного исследования разобраться в предложениях, связанных с реализацией алгоритмической прозрачности в области правового регулирования обработки данных.

1. Понятие алгоритмической прозрачности

В Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 г.[2] проблема алгоритмической прозрачности систем искусственного интеллекта отнесена к концептуальным проблемным направлениям регулирования отношений в сфере технологий искусственного интеллекта и робототехники[3].

«Алгоритмическая прозрачность» информационной модели позволяет получить представление о логике функционирования информационной модели, реализуемой

---

1 Указ Президента Российской Федерации № 490 от 10.11.2019 (вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 года). (2019). Собрание законодательства Российской Федерации, 41, ст. 5700.

2 Распоряжение Правительства Российской Федерации № 2129-р от 19.08.2020. (2020). Собрание законодательства Российской Федерации, 35, ст. 5593.

3 Там же.

ИИ при заданных входных данных. Однако необходимо сразу оговориться, что сложность алгоритмов ИИ не позволяет описать любой алгоритм ИИ так, чтобы его логика стала понятна обывателю. Наиболее простой с точки зрения восприятия происходящих алгоритмических этапов является линейная модель ИИ и ее математическая интерпретация. Наиболее сложной для понимания логики функционирования модели ИИ является глубокая архитектура ИИ[4].

Иными словами, алгоритмическая прозрачность – это объяснимость работы «ИИ 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»[5]. С точки зрения некоторых исследователей, технологические разработки в области ИИ и алгоритмов стали неотъемлемой частью государственного управления (Feijóo et al., 2020; Carlsson & Rönnblom, 2022; Balasubramaniam et al., 2023; Green, 2022).

2. Защитит ли пользовательские данные понимание логики работы алгоритма обработки данных?

В 1999 г. Л. Лессиг был одним из первых авторов, который отдавая должное правовым и социальным нормам в обеспечении правового регулирования отношений, возникающих в ИКТ-сфере, признал программный код равнозначной составляющей в области регулирования информационных отношений, именно он определяет архитектуру пространства ИКТ-сферы. Программный код позволяет достичь наилучшего результата в регулировании отношений, возникающих в информационной сфере (Lessig, 1999).

Программный код формализует логику работы алгоритма. Все чаще звучат предложения об обеспечении алгоритмической прозрачности программного обеспечения. «Прозрачность алгоритмов становится своеобразной формой контроля, а прозрачность алгоритмического принятия решений служит тому, чтобы несправедливые дискриминации могли быть обнаружены и оспорены» (Талапина, 2020). Хотя есть противоположная точка зрения. Так, некоторые ученые считают, что требование об алгоритмической прозрачности – это действия, направленные на взаимодействие разработчиков ИТ с органами власти в целях нормализации поведения граждан (Wang, 2022).

По нашему мнению, требование о прозрачности алгоритмов не принесет желаемого результата, поскольку разобраться в логике алгоритма ИИ не всегда под силу даже специалисту. В связи с этим несправедливые дискриминации, заложенные в алгоритмы ИИ, не всегда могут быть обнаружены и оспорены.

Анализируя тенденцию реализации алгоритмической прозрачности, попытаемся привести позиции за и против раскрытия логики алгоритма.

---

4 Корешкова, Т. (2020, 29 декабря). Объяснимый искусственный интеллект. Научно-технический центр ФГУП «ГРЧЦ». https://clck.ru/36h6w6

5 Указ Президента РФ № 490 от 10.11.2019 (вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 года). (2019). Собрание законодательства Российской Федерации, 41, ст. 5700.

Так, обыватель, возлагая надежду на раскрытие логики алгоритма, считает, что это позволит ему понять его логику, однако такие предложения, хотя и не безосновательны, но имеют свои недостатки. Во-первых, большинству людей, не имеющих компетенций в области программирования и разработки информационных технологий, раскрытие логики работы алгоритма не даст никакой информации. Во-вторых, если пользователь алгоритма, например, ИИ, разберется в его логике, то внесение изменений в алгоритм будет невозможно в связи с тем, что для этого потребуется пересмотреть весь заложенный в алгоритм математический инструментарий. В-третьих, раскрытие логики работы алгоритма не должно противоречить требованиям законодательства об интеллектуальной собственности, поскольку интеллектуальные права на алгоритмы принадлежат его правообладателям, разработчикам. Соответственно, раскрытие логики работы алгоритма может происходить в строго ограниченных случаях.

Существуют и иные позиции ученых, которые предлагают заменить раскрытие логики работы алгоритмов страхованием рисков, связанных с обеспечением информационной безопасности. Например, пользователь, садясь в самолет и доверяя свою жизнь перевозчику, предварительно не изучает логику работы программного обеспечения самолета. Все риски берут на себя перевозчик, страховщик и другие лица, ответственные за перевозку пассажиров (Остроумов, 2015). Разве в случае обработки данных мы не можем использовать эту же правовую схему регулирования отношений?

«Перевозчиками» пользовательских данных являются различные информационные посредники, например, провайдеры, операторы обработки персональных данных. Учитывая существующую высокую вероятность того, что раскрытие логики работы алгоритма фактически ничего не даст пользователю, то не будет ли более эффективным применение системы страхования к отношениям в ИКТ-сфере, как в случае с воздушными перевозками? В этом случае будут застрахованы риски «гибели» пользовательских данных или несанкционированного доступа к ним, а также ответственность информационных посредников или операторов персональных данных.

Однако и в этом случае есть свои подводные камни. Так, в случае с воздушными перевозками все этапы, начиная от создания самолета до его полета, жестко регламентируются нормами правового и технического регулирования, чего не происходит в случае создания алгоритмов, информационных моделей и их использования. Стандартизация информационных технологий в целях обеспечения информационной безопасности является добровольной. Обязательной стандартизации подлежат такие информационные системы, как критическая информационная инфраструктура Российской Федерации и системы, обрабатывающие персональные данные.

В связи с этим проведение аналогии между обеспечением безопасности пользователя ИТ через обеспечение алгоритмической прозрачности и безопасностью воздушных перевозок возможно только в условиях жесткой регламентации создания ИТ и их использования нормами правового и технического регулирования.

Вопрос доверия в области информационной безопасности волнует умы ученых разных государств (Cui et al., 2022; Bujold et al., 2022; Zhu et al., 2023). Доверие определяется как культурная ценность, которая иногда может вступать в противоречие с национальной политикой в области ИИ (Robinson, 2020; Xu et al., 2022; Li, 2022).

3. Сравнительный анализ принципов алгоритмической прозрачности

Эксперты Сообщества справедливости, подотчетности и прозрачности в области машинного обучения (Fairness, Accountability, and Transparency in Machine Learning – FAT) определяют пять принципов алгоритмической прозрачности: справедливость, проверяемость, объяснимость, ответственность разработчиков и точность работы[6]. Как видим, в предложениях Сообщества справедливости, подотчетности и прозрачности в области машинного обучения объяснимость логики работы ИИ стоит на третьем месте. Скорее всего, это связано именно с тем, что алгоритмическая открытость не во всех случаях может решить вопросы обеспечения безопасности пользователя ИТ.

Учеными предлагается дополнить эти пять принципов принципом внесения изменений в логику работы алгоритма ИИ в случае несогласия с ее функционированием (Малышкин, 2019; Gordon et al., 2022). Однако такие предложения вызывают у нас опасения, поскольку в этом случае возможно нарушение законодательства об интеллектуальной собственности. Большинство компаний не стремятся раскрывать свои алгоритмы и делать их прозрачными, «ссылаясь на потенциальные игры пользователей, которые могут негативно повлиять на предсказательную способность алгоритма» (Qiaochu et al., 2020; Stahl et al., 2022).

Со своей стороны хотели бы подчеркнуть, что отсутствие в сформулированных пяти принципах «возможности изменения логики работы ИИ» объяснимо. Такие алгоритмы разрабатываются коллективом программистов, изменение логики работы одной части алгоритма ИИ сделает неработоспособной математическую модель всего алгоритма (Varsha, 2023; Lang & Shan, 2000; Akter et al., 2022). Если бы человеческий мозг мог решить задачу обработки больших, неструктурированных данных, то алгоритмы ИИ были бы бесполезны.

В Российской Федерации в соответствии с Национальной стратегией развития искусственного интеллекта на период до 2030 г. принципами развития и использования ИИ, соблюдение которых обязательно, являются: защита прав и свобод человека, технологический суверенитет, целостность инновационного цикла, разумная бережливость, поддержка конкуренции, безопасность и прозрачность[7].

Под безопасностью понимается «недопустимость использования искусственного интеллекта в целях умышленного причинения вреда гражданам и юридическим лицам, а также предупреждение и минимизация рисков возникновения негативных последствий использования технологий искусственного интеллекта»[8]. Прозрачность определена 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» (п. 19)[9].

---

6 Principles for Accountable Algorithms and a Social Impact Statement for Algorithms. https://clck.ru/36h7GL

7 Указ Президента РФ № 490 от 10.11.2019 (вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 г.). (2019). Собрание законодательства Российской Федерации, 41, ст. 5700.

8 Там же.

9 Там же.

Подчеркнем, что, если Сообществом справедливости, подотчетности и прозрачности в области машинного обучения алгоритмическая прозрачность определена через пять принципов, то российское законодательство алгоритмическую прозрачность включило в принципы развития и использования ИИ. Принцип недискриминационного доступа пользователей ИИ к информации о применяемых алгоритмах работы ИИ пересекается с принципами, сформулированными Сообществом справедливости, подотчетности и прозрачности в области машинного обучения.

Не содержит законодательство Российской Федерации также требования об опубликовании правил, определяющих основные алгоритмические обработки пользовательских данных, в отличие от французского законодательства. Так, в соответствии с Законом Франции «О цифровой республике» от 7 октября 2016 г. такие правила должны публиковаться на сайте государственного органа (Талапина, 2020).

В Евросоюзе безопасность персональных данных резидентов Евросоюза (Su et al., 2023), а также прозрачность их обработки алгоритмами (Matheus et al., 2021; Kempeneer, 2021), в том числе ИИ (Kempeneer et al., 2023; de Bruijn et al., 2022), регулируются Регламентом защиты персональных данных (Stöger et al., 2021) – General Data Protection Regulation (далее – GDPR), вступившим в силу в 2018 г. (Mourby et al., 2021).

4. Разработка и принятие Общего регламента по защите данных

Несмотря на то, что защита частной жизни напрямую не связана с защитой персональных данных, их использование позволяет идентифицировать личность человека и соответственно получить информацию о его частной жизни. Таким образом, надежно защищенные персональные данные уменьшают риски получения информации о частной жизни лица (Bolton et al., 2021; Leerssen, 2023). Необходимость борьбы с различными нарушениями законодательства в области персональных данных и частной жизни лица, а также минимизация последствий таких деяний, способствовали разработке и принятию GDPR (Willems et al., 2022; Custers & Heijne, 2022).

Одним из первых дел, связанных с незаконным получением информации о частной жизни лица путем незаконного доступа к базе персональных данных, было дело Uber. Взломы базы персональных данных Uber произошли в 2014 и 2016 гг., что позволило злоумышленникам отслеживать местонахождение каждого пользователя Uber в режиме реального времени. В 2017 г. Федеральная торговая комиссия (далее – ФТС) обвинила Uber в отсутствии должного контроля доступа сотрудников к базам данных пользователей и водителей системы Uber, а также в нарушении системы обеспечения безопасности информации. В дальнейшем между Uber и ФТС было подписано соглашение, в рамках которого Uber обязалась проводить сторонние проверки в течение двадцати лет и реализовать программу защиты конфиденциальности[10].

В 2017 г. ФТС включила в соглашение дополнительные положения, обязывающие Uber проводить аудит их системы и представлять отчеты в ФТС, а также раскрывать информацию о вознаграждениях и условиях соглашений, которые Uber заключает с третьими лицами, осуществляющими мониторинг уязвимостей программного обеспечения Uber. В соответствии с последней редакцией соглашения Uber:

[10] Uber criminal complaint raises the stakes for breach response. https://clck.ru/37AXba

– может быть привлечена к гражданско-правовым санкциям, если она не уведомит ФТС об инцидентах, связанных с несанкционированным доступом к информации о пользователях и водителях Uber;

– запрещено искажать информацию об уровне защиты информации в системе, об условиях обеспечения конфиденциальности, безопасности и целостности личной информации, а также о том, как она контролирует внутренний доступ к личной информации потребителей;

– должна внедрить комплексную программу конфиденциальности и в течение 20 лет получать раз в два года независимые оценки обеспечения безопасности своей информационной системы, проводимые третьими лицами. Компания Uber должна предоставлять эти оценки в ФТС и подтверждать выполнение принятой программы конфиденциальности, которая должна содержать условия обеспечения безопасности, имеющиеся в соглашении с ФТС;

– обязана хранить информацию о местоположении пользователя в системе, защищенную паролем и шифром;

– обязана проводить ежегодное обучение сотрудников, ответственных за обработку личной информации, методам защиты данных и обеспечения безопасности, принятой в Uber, а также использовать новейшие методы контроля безопасности;

– обязана использовать передовые методы защиты данных для защиты персональных данных водителей;

– должна назначить одного или нескольких сотрудников для координации и контроля программы обеспечения безопасности и конфиденциальности, а также проводить регулярные оценки эффективности внутреннего контроля и процедур Uber, связанных с защитой личной информации и информации о географическом местоположении своих сотрудников и клиентов Uber;

– обязана использовать многофакторную аутентификацию, прежде чем любой сотрудник сможет получить доступ к конфиденциальной личной информации клиента, а также другие надежные методы обеспечения безопасности данных[11].

В связи с нарушениями, произошедшими в 2014 и 2016 гг., компания Uber заплатила штраф в размере 148 млн долларов[12]. 20 августа 2020 г. в отношении бывшего начальника службы безопасности Uber было возбуждено уголовное дело, в котором было предъявлено обвинение в воспрепятствовании правосудию и предполагаемой попытки сокрытия утечки данных, произошедшей в 2016 г.

5. Реализация общего регламента по защите данных

GDPR определяет право каждого на защиту своих персональных данных в соответствии с ч. 1 ст. 16 Договора о функционировании Европейского союза (TFEU)[13] и ч. 1

[11] Там же.

[12] Uber to Pay $148 Million Fine for Massive Data Breach That Exposed 57 Million Users' Personal Info. https://clck.ru/36h7RG

[13] Договор о функционировании Европейского союза [рус., англ.] (вместе со «Списком, предусмотренным в статье 38...», «Заморскими странами и территориями, к которым применяются положения части четвертой Договора...») (подписан в г. Риме 25.03.1957) (с изм. и доп. от 13.12.2007). СПС «КонсультантПлюс».

ст. 8 Хартии Европейского союза об основных правах[14] (далее – Хартия)[15], а также «право на неприкосновенность частной жизни» (ст. 7 Хартии)[16].

GDPR требует от компаний, обрабатывающих персональные данные резидентов Евросоюза или осуществляющих свою деятельность на территории государства Евросоюза, соблюдения не только правовых требований, но и организационно-технических, которые должны быть учтены разработчиками еще на этапе проектирования информационных технологий, и названных как «конфиденциальность при проектировании» (ст. 3 GDPR). Требования об обеспечении конфиденциальности в цифровом мире посредством «защиты персональных данных при проектировании и по умолчанию» утверждены European Data Protection Board (EDPB) в Руководстве 4/2019 «Встроенная защита данных и по умолчанию» (Guidelines 4/2019 on Article 25 Data Protection by Design and by Default)[17].

Реализация данных требований вызвала у компаний множество вопросов о процедурах их реализации. В связи с чем Европейский совет по защите данных – EDPB – и Европейский надзорный орган по защите данных (European Data Protection Supervisor – EDPS) представили разъяснения. В Заявлении EDPB 03/2021 «Положение о конфиденциальности», принятом Советом по защите частной жизни и конфиденциальности при использовании услуг электронной связи, указывается, что предлагаемое Положение ни при каких обстоятельствах не должно снижать уровень защиты, определенный в действующей Директиве 2002/58/EC[18]. Положение должно дополнять GDPR, предоставляя дополнительные надежные гарантии конфиденциальности и защиту всех типов электронных сообщений[19]. Директива 2002/58/EC охватывает обработку персональных данных и защиту конфиденциальности, включая требования об обеспечении безопасности сетей и услуг; конфиденциальности общения; доступа к сохраненным данным; обработке данных о трафике и местоположении; идентификации; общедоступных справочниках подписчиков, а также запрете коммерческих сообщений (спама)[20]. EDPB особое внимание уделяет безопасности персональных данных, обрабатываемых работодателями. EDPB четко определяет случаи и условия, при которых работодатели могут получить доступ к персональным данным сотрудников, а также ответственность за чрезмерный сбор данных с помощью технологий анализа и обработки данных. Например, применение работодателем систем геолокации, технологий постоянного контроля перемещений и поведения сотрудника.

---

14 Хартия Европейского союза об основных правах (2007/C 303/01) [рус., англ.] (Вместе с «Разъяснениями...» (2007/C 303/02)) (Принята в г. Страсбурге 12.12.2007). СПС «КонсультантПлюс».

15 Там же.

16 Charter of Fundamental Rights of the European Union. https://clck.ru/36h7Tn

17 Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0 Adopted on 20 October 2020. https://clck.ru/36h7Ug

18 E-privacy Directive 2009/136/EC. https://clck.ru/36h7VG

19 Statement 03/2021 on the ePrivacy Regulation Adopted on 9 March 2021. https://clck.ru/36h7WF

20 E-privacy Directive 2009/136/EC. https://clck.ru/36h7VG

Поскольку функции EDPB и EDPS пересекаются, то для разграничения их деятельности был принят Меморандум о взаимопонимании между EDPB и EDPS[21], в соответствии с которым EDPB обеспечивает единство правоприменительной практики GDPR, а EDPS обеспечивает единство подходов национальных надзорных органов. При этом EDPB и EDPS могут издавать совместные документы по вопросам защиты персональных данных.

6. Концепция конфиденциальности при проектировании информационных технологий

Концепция конфиденциальности при проектировании была разработана задолго до принятия GDPR. Так, в 1995 г. в Директиву 95/46 / ЕС «О защите данных»[22] было включено положение о том, что «в целях защиты безопасности данных технические и организационные меры должны быть определены и приняты еще на стадии планирования системы обработки данных» (ст. 46 Директивы 95/46 / ЕС).

22 июня 2011 г. EDPS вынес на обсуждение концепцию изменения подхода к регулированию защиты персональных данных и обеспечению их конфиденциальности[23] в форме публичного мнения этой организации. В целях необходимости и целесообразности учета требований защиты персональных данных посредством конфиденциальности при проектировании ученые предложили изменить концепцию защиты персональных данных, изложив ее в следующих семи принципах[24]:

1. Меры обеспечения конфиденциальности при проектировании должны быть превентивными и учитывать возможные риски и угрозы, а не являться ответной реакцией на нарушения конфиденциальности.

2. Решение проблемы обеспечения конфиденциальности в информационных системах должно быть реализовано в системе еще на уровне ее разработки, а не являться опцией для пользователя.

3. Возможные риски и угрозы должны учитываться еще на этапе проектирования технологии, а также быть прописаны в стандартах информационной безопасности и учитывать контекст данных. Методы обеспечения безопасности персональных данных должны постоянно обновляться.

4. Конфиденциальность должна быть реализована на всех этапах жизненного цикла обработки персональных данных, это позволит обеспечить непрерывность управления безопасностью. Применяемые стандарты безопасности должны гарантировать конфиденциальность, целостность и доступность личных данных на протяжении всего их жизненного цикла, а также реализацию методов безопасного уничтожения, шифрования, контроля доступа и ведения журналов.

5. Должны обеспечиваться мониторинг, оценка и проверка соблюдения политик и процедур конфиденциальности, открытость и прозрачность в целях соблюдения

---

21 Memorandum of Understanding. https://clck.ru/36h7ic

22 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. https://clck.ru/36h7jM

23 The History of the General Data Protection Regulation. https://clck.ru/36h7jw

24 The Seven Principles. https://goo.su/mn8ob7

принципа подотчетности и создания условий для обеспечения доверия со стороны субъектов персональных данных и контрагентов, а также унификации деловой практики. Физическим лицам должны быть доступны информация о политиках и методах управления личной информацией, механизмы соблюдения требований и рассмотрения жалоб.

6. Не должен стоять выбор между безопасностью и функциональностью.

7. Права и интересы субъектов персональных данных должны быть основой для проектирования конфиденциальности.

Данные принципы были одними из первых, которые учли фактически все возможные риски нарушения обработки персональных данных. Однако предлагались и другие концепции[25].

7. Понятия «персональные данные» и «конфиденциальность» в соответствии с законодательством Европейского союза

В соответствии с GDPR под персональными данными понимается любая информация, относящаяся к физическому лицу, которое можно идентифицировать. В соответствии со ст. 4 GDPR к таким данным можно, например, отнести ссылку на идентификатор, такой как имя, идентификационный номер, данные о местоположении, онлайн-идентификатор, любые факторы, специфичные для физических, физиологических, генетических, ментальных, экономических, культурных или социальных идентичностей этого физического лица.

В тексте GDPR определение понятия «конфиденциальность» отсутствует, однако дается отсылка к Директиве 2002/58/EC Европейского парламента и Совета от 12 июля 2002 г. относительно обработки персональных данных и защиты конфиденциальности в секторе электронных коммуникаций (далее – Директива 2002/58/EC)[26]. GDPR использует понятие «персональные данные». По мнению EDPB[27], EDPS[28], а также ENISA[29] в контексте проектируемой конфиденциальности понятия «персональные данные» и «конфиденциальность» следует рассматривать как синонимы. Кроме того, руководящие рекомендации EDPB, EDPS, ENISA определяют, что для ситуаций, не имеющих особого значения, не должно делаться различий между защитой персональных данных и обеспечением конфиденциальности при проектировании и по умолчанию.

---

25 Langheinrich, M. Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems: Distributed Systems Group Institute of Information Systems, IFW Swiss Federal Institute of Technology, ETH Zurich 8092 Zurich, Switzerland. https://clck.ru/36h7qq

26 Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) (OJ L 201, 31.7.2002, p. 37).

27 EDPB – это независимый орган Европейского союза, созданный и функционирующий на основании GDPR. EDPB способствует обеспечению согласованного применения GDPR, для чего обладает рядом полномочий, установленных ст. 70 GDPR. В частности, в полномочиях этого органа находится издание руководства, рекомендаций и лучших практик применения GDPR.

28 EDPS – это независимый орган ЕС, контролирующий деятельность национальных надзорных органов, которые должны быть созданы в соответствии с разделом V GDPR.

29 Агентство Европейского союза по кибербезопасности ENISA. https://clck.ru/N598K

Однако такая общность в понимании защиты персональных данных и обеспечения конфиденциальности при проектировании и по умолчанию принимается не для всех случаев. В опубликованном мнении EDPS[30] «О защите конфиденциальности при проектировании и по умолчанию» (opinion on privacy by design and by default) делается различие между двумя понятиями – «конфиденциальность при проектировании» и «защита данных при проектировании» (privacy by design и data protection by design). Понятие «конфиденциальность при проектировании» используется для обозначения системы технологических мер, направленных на обеспечение конфиденциальности, выработанной в ходе международных дебатов за последние несколько десятилетий. Данное понятие определяет правовой режим информации, заключающийся в ограничении доступа к ней, и означает «защиту данных с помощью технологического проектирования»[31] (Zharova, 2020).

Под «защитой данных при проектировании» понимается предварительное решение вопросов защиты данных и конфиденциальности на этапе проектирования технологии для всех действий пользователя[32] (Zharova, 2019).

Обсуждения степени различия данных терминов продолжаются и до настоящего времени. Так, разработчики разъяснений о применении GDPR[33] пишут, что все еще существует неопределенность в отношении того, что означает конфиденциальность при проектировании и как ее можно реализовать. Такая проблема возникает в связи с тем, что, с одной стороны, Директива 95/46/EC (Directive 95/46/EC)[34] в некоторых государствах-членах реализуется не в полной мере. С другой – принцип конфиденциальности при проектировании, содержащийся в GDPR, определяет, что руководящие принципы безопасности данных требуют, чтобы организационные и технические меры были приняты еще на уровне планирования информационной системы. Например, принцип GDPR обеспечения целостности и конфиденциальности определяет необходимость защиты данных от несанкционированного доступа или их незаконной обработки, а также от случайной потери, уничтожения или повреждения[35]. Однако законодательство Евросоюза оставляет полностью открытым вопрос о принимаемых ответственными лицами защитных мерах. Например, достаточно ли анонимизации имени человека для реализации требований законодательства?

GDPR предлагает использовать шифрование или анонимизацию данных в качестве возможной меры обеспечения конфиденциальности при проектировании. Однако это предложение не позволяет понять, как эта мера в дальнейшем будет

30 Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. https://clck.ru/36h7v3

31 GDPR Privacy by Design. https://clck.ru/36h7vZ

32 Data protection by design and default. https://goo.su/Hxoh2d

33 GDPR Privacy by Design. https://clck.ru/36h7vZ

34 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. https://clck.ru/36h7jM

35 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) of 27 April 2016. https://clck.ru/34U2FN

согласовываться с требованием GDPR об аутентификации пользователя и технической реализацией права на возражение.

В итоге авторы разъяснений о применении GDPR определили термин «конфиденциальность при проектировании» как «защиту данных с помощью технологического проектирования», они считают, что «в процедурах обработки данных лучше всего придерживаться защиты данных, когда она уже интегрирована в технологию на этапе ее проектирования»[36].

Заключение

Проблема защиты персональных данных, контроль принципов обработки больших пользовательских данных, защита частной жизни лица с каждым годом только обостряются. Необходимость обеспечения защиты персональных данных и частной жизни лица как пользователя ИТ ставит перед законодателем задачу обеспечения открытости логики работы алгоритмов информационных технологий. Например, для достижения этой цели в GDPR определено требование о реализации проектируемой конфиденциальности при разработке ИТ, которое было предложено еще в 1995 г. Требование о реализации принципа алгоритмической прозрачности в системах ИИ было предложено намного позже, в 2019 г., российскими законодателями, а зарубежными законодателями – в 2018 г.

Алгоритмы обработки данных только усложняются, в связи с чем все чаще звучат законодательные предложения о раскрытии логики их работы, например в системах ИИ; нужно понимать, что такие предложения можно реализовать не для всех алгоритмов. Объяснить сложный математический инструментарий простыми словами, которые будут понятны каждому обывателю, вряд ли удастся.

Однако это не значит, что не существует выхода из этой сложной технико-правовой ситуации. Мы считаем, что развитие области стандартизации информационной безопасности, включение в правовые акты требования о разработке ИТ, соответствующих требованиям стандартизации, позволит минимизировать риски, связанные с неправомерной обработкой больших пользовательских данных и получением информации о частной жизни лица.

Список литературы

Гулемин, А. Н. (2022). Пределы обработки больших объемов данных для целей получения информации о человеке: правовой аспект. *Электронное приложение к Российскому юридическому журналу*, *6*, 52–57. http://doi.org/10.34076/22196838_2022_6_52

Кутейников, Д. Л., Ижаев, О. А., Зенин, С. С., Лебедев, В. А. (2020). Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика». *Lex russica (Русский закон)*, *73*(6), 146. https://doi.org/10.17803/1729-5920.2020.163.6.139-148

Малышкин, А. В. (2019). Интегрирование искусственного интеллекта в общественную жизнь: некоторые этические и правовые проблемы. Вестник Санкт-Петербургского университета. *Право*, *10*(3), 444–460. https://doi.org/10.21638/spbu14.2019.303

Остроумов, Н. Н. (2015). *Правовой режим международных воздушных перевозок*. Москва: Статут. https://elibrary.ru/ulcfpl

[36] GDPR Privacy by Design. https://clck.ru/36h7vZ

Талапина, Э. В. (2020). Алгоритмы и искусственный интеллект сквозь призму прав человека. *Журнал российского права*, *10*, 25–39. https://doi.org/10.12737/jrl.2020.118

Akter, Sh., Dwivedi, Y. K., Sajib, Sh., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, *159*, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Bolton, M., Raven, R., & Mintrom, M. (2021). Can AI transform public decision-making for sustainable development? An exploration of critical earth system governance questions. *Earth System Governance*, *9*, 100116. https://doi.org/10.1016/j.esg.2021.100116

Bujold, A., Parent-Rocheleau, X., & Gaudet, M.-C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management, justice perception, and intention to quit among truck drivers. *Computers in Human Behavior Reports*, *8*, 100245. https://doi.org/10.1016/j.chbr.2022.100245

Carlsson, V., & Rönnblom, M. (2022). From politics to ethics: Transformations in EU policies on digital technology. *Technology in Society*, *71*, 102145. https://doi.org/10.1016/j.techsoc.2022.102145

Cui, M., Mariani, M. S., & Medo, M. (2022). Algorithmic bias amplification via temporal effects: The case of PageRank in evolving networks. *Communications in Nonlinear Science and Numerical Simulation*, *104*, 106029. https://doi.org/10.1016/j.cnsns.2021.106029

Custers, B., & Heijne, A.-S. (2022). The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice. *Computer Law & Security Review*, *46*, 105727. https://doi.org/10.1016/j.clsr.2022.105727

de Bruijn, H., Warnier, M., & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 101666. https://doi.org/10.1016/j.giq.2021.101666

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., Potgieter, P., Vu, K., Whalley, J., & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44*(6), 101988. https://doi.org/10.1016/j.telpol.2020.101988

Gordon, G., Rieder, B., & Sileno, G. (2022). On mapping values in AI Governance. *Computer Law & Security Review*, *46*, 105712. https://doi.org/10.1016/j.clsr.2022.105712

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, *45*, 105681. https://doi.org/10.1016/j.clsr.2022.105681

Kempeneer, Sh. (2021). A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. *Government Information Quarterly*, *38*(3), 101578. https://doi.org/10.1016/j.giq.2021.101578

Kempeneer, Sh., Pirannejad, A., & Wolswinkel, J. (2023). Open government data from a legal perspective: An AI-driven systematic literature review. *Government Information Quarterly*, 101823. https://doi.org/10.1016/j.giq.2023.101823

Lang, H., & Shan, C. (2000). Bias phenomenon and compensation in multiple target tracking algorithms. *Mathematical and Computer Modelling*, *31*(8–9), 147–165. https://doi.org/10.1016/S0895-7177(00)00063-7

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review*, *48*, 105790. https://doi.org/10.1016/j.clsr.2023.105790

Lessig, L. (1999). *Code and other laws of cyberspace*. New York: Basic Books.

Li, Z. (2022). Affinity-based algorithmic pricing: A dilemma for EU data protection law. *Computer Law & Security Review*, *46*, 105705. https://doi.org/10.1016/j.clsr.2022.105705

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, *38*(1), 101550. https://doi.org/10.1016/j.giq.2020.101550

Mourby, M., Ó Cathaoir, K., & Bjerre Collin, C. (2021). Transparency of machine-learning in healthcare: The GDPR & European health law. *Computer Law & Security Review*, *43*, 105611. https://doi.org/10.1016/j.clsr.2021.105611

Qiaochu, W., Yan, H., Stefanus, J., & Param Vir, S. (2020, July 15). *Algorithmic Transparency with Strategic Users*. http://dx.doi.org/10.2139/ssrn.3652656

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, *63*, 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Stahl, B. C., Rodrigues, R., Santiago, N., & Macnish, K. (2022). A European Agency for Artificial Intelligence:

Protecting fundamental rights and ethical values. *Computer Law & Security Review*, *45*, 105661. https://doi.org/10.1016/j.clsr.2022.105661

Stefik, M. (1996). Letting loose the light: Igniting commerce in electronic publication. In M. Stefik (Ed.), *Internet dreams: Archetypes, myths, and metaphors* (pp. 219–253). Cambridge, MA: MIT Press.

Stöger, K., Schneeberger, D., Kieseberg, P., & Holzinger, A. (2021). Legal aspects of data cleansing in medical AI. *Computer Law & Security Review*, *42*, 105587. https://doi.org/10.1016/j.clsr.2021.105587

Su, Zh., Bentley, B. L., McDonnell, D., Cheshmehzangi, A., Ahmad, J., Šegalo, S., Pereira da Veiga, C., & Xiang, Yu-Tao. (2023). China's algorithmic regulations: Public-facing communication is needed. *Health Policy and Technology*, *12*(1), 100719. https://doi.org/10.1016/j.hlpt.2022.100719

Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, *3*(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Wang, H. (2022). Transparency as Manipulation? Uncovering the Disciplinary Power of Algorithmic Transparency. *Philosophy & Technology*, *35*, 69. https://doi.org/10.1007/s13347-022-00564-w

Willems, J., Schmidthuber, L., Vogel, D., Ebinger, F., & Vanderelst, D. (2022). Ethics of robotized public services: The role of robot design and its actions. *Government Information Quarterly*, *39*(2), 101683. https://doi.org/10.1016/j.giq.2022.101683

Xu, J., Xiao, Yu., Wang, W. Hu., Ning, Yu., Shenkman, E. A., Bian, J., & Wang, F. (2022). Algorithmic fairness in computational medicine. *eBioMedicine*, *84*, 104250. https://doi.org/10.1016/j.ebiom.2022.104250

Zharova, A. (2019). Ensuring the information security of information communication technology users in Russia. *International Journal of Cyber Criminology*, *13*(2), 255–269. EDN: https://elibrary.ru/ltmesv. DOI: https://doi.org/10.5281/zenodo.3698141

Zharova, A. (2020). The protect mobile user data in Russia. *International Journal of Electrical and Computer Engineering*, *10*(3), 3184–3192. EDN: https://www.elibrary.ru/juzboh. DOI: https://doi.org/10.11591/ijece.v10i3.pp3184-3192

Zhu, H., Sallnäs Pysander, E.-L., & Söderberg, I.-L. (2023). Not transparent and incomprehensible: A qualitative user study of an AI-empowered financial advisory system. *Data and Information Management*, 100041. https://doi.org/10.1016/j.dim.2023.100041

Сведения об авторе

Жарова Анна Константиновна – доктор юридических наук, доцент, старший научный сотрудник, Институт государства и права Российской академии наук
Адрес: 420100, Российская Федерация, г. Москва, ул. Знаменка, 10
E-mail: anna_jarova@mail.ru
ORCID ID: https://orcid.org/0000-0002-2981-3369
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56964137900
WoS Researcher ID: https://www.webofscience.com/wos/author/record/H-4012-2015
Google Scholar ID: https://scholar.google.com/citations?user=g8ij3BsAAAAJ
РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=151076

Конфликт интересов

Автор является главным редактором журнала, статья прошла рецензирование на общих основаниях.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.19.61 / Правовое регулирование информационной безопасности
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 22 мая 2023 г.
Дата одобрения после рецензирования – 21 августа 2023 г.
Дата принятия к опубликованию – 30 ноября 2023 г.
Дата онлайн-размещения – 15 декабря 2023 г.

Research article
DOI: https://doi.org/10.21202/jdtl.2023.42

Achieving Algorithmic Transparency and Managing Risks of Data Security when Making Decisions without Human Interference: Legal Approaches

Anna K. Zharova

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

Keywords

algorithmic transparency,
artificial intelligence,
confidentiality,
data protection,
data security,
digital technologies,
GDPR,
information technologies,
law,
personal data

Abstract

Objective: to compare modern approaches in law to the use of program codes and algorithms in decision-making that meet the principles of transparency and openness, as well as the increasingly stringent requirements for ensuring the security of personal and other big data obtained and processed algorithmically.

Methods: the main methods for researching the principle of transparency in algorithmic decision-making were formal-legal and comparative analysis of legal acts and international standards of information security, as well as the principles and legal constructions contained in them.

Results: it was determined that the development of information security standardization, inclusion in legal acts of requirements for the development of information technologies that comply with the principles of transparency and openness of applied algorithms will minimize the risks associated with the unlawful processing of users' big data and obtaining information about their privacy. Proposals were identified, related to the implementation of algorithmic transparency in the field of data processing legal regulation. Recommendations were formulated, based on which the legislator can solve the problem of ensuring the openness of the logic of information technology algorithms with regard to modern standards of information security.

Scientific novelty: it consists in the substantiation of new trends and relevant legal approaches, which allow revealing the logic of data processing by digital and information technologies, based on the characterization of European standards of the "privacy by design" concept in new digital and information technologies of decision-making and data protection, as well as on the new legal requirements for artificial intelligence systems, including the requirement to ensure algorithmic transparency, and criteria for personal data and users' big data processing. This said, data protection is understood as a system of legal, technical and organizational principles aimed at ensuring personal data confidentiality.

Practical significance: it is due to the need to study the best Russian and international practices in protecting the privacy of users of digital and information technologies, as well as the need for legislative provision of requirements for the use of algorithms that meet the principles of transparency and openness of personal data processing, taking into account the need to ensure confidentiality at all stages of the life cycle of their processing, which will ensure the continuity of security management.

For citation

Zharova, A. K. (2023). Achieving Algorithmic Transparency and Managing Risks of Data Security when Making Decisions without Human Interference: Legal Approaches. *Journal of Digital Technologies and Law, 1*(4), 973–993. https://doi.org/10.21202/jdtl.2023.42

References

Akter, Sh., Dwivedi, Y. K., Sajib, Sh., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, *159*, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Bolton, M., Raven, R., & Mintrom, M. (2021). Can AI transform public decision-making for sustainable development? An exploration of critical earth system governance questions. *Earth System Governance*, *9*, 100116. https://doi.org/10.1016/j.esg.2021.100116

Bujold, A., Parent-Rocheleau, X., & Gaudet, M.-C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management, justice perception, and intention to quit among truck drivers. *Computers in Human Behavior Reports*, *8*, 100245. https://doi.org/10.1016/j.chbr.2022.100245

Carlsson, V., & Rönnblom, M. (2022). From politics to ethics: Transformations in EU policies on digital technology. *Technology in Society*, *71*, 102145. https://doi.org/10.1016/j.techsoc.2022.102145

Cui, M., Mariani, M. S., & Medo, M. (2022). Algorithmic bias amplification via temporal effects: The case of PageRank in evolving networks. *Communications in Nonlinear Science and Numerical Simulation*, *104*, 106029. https://doi.org/10.1016/j.cnsns.2021.106029

Custers, B., & Heijne, A.-S. (2022). The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice. *Computer Law & Security Review*, *46*, 105727. https://doi.org/10.1016/j.clsr.2022.105727

de Bruijn, H., Warnier, M., & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 101666. https://doi.org/10.1016/j.giq.2021.101666

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., Potgieter, P., Vu, K., Whalley, J., & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44*(6), 101988. https://doi.org/10.1016/j.telpol.2020.101988

Gordon, G., Rieder, B., & Sileno, G. (2022). On mapping values in AI Governance. *Computer Law & Security Review*, *46*, 105712. https://doi.org/10.1016/j.clsr.2022.105712

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, *45*, 105681. https://doi.org/10.1016/j.clsr.2022.105681

Gulemin, A. (2022). Limits of big data processing for the purposes of obtaining information about a person: a legal aspect. In Elektronnoe prilozhenie k "Rossiiskomu yuridicheskomu zhurnalu", 6, 52–57. (In Russ.). http://doi.org/10.34076/22196838_2022_6_52

Kempeneer, Sh. (2021). A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. *Government Information Quarterly*, *38*(3), 101578. https://doi.org/10.1016/j.giq.2021.101578

Kempeneer, Sh., Pirannejad, A., & Wolswinkel, J. (2023). Open government data from a legal perspective: An AI-driven systematic literature review. *Government Information Quarterly*, 101823. https://doi.org/10.1016/j.giq.2023.101823

Kuteynikov, D. L., Izhaev, O. A., Zenin, S. S., & Lebedev, V. A. (2020). Algorithmic transparency and accountability: legal approaches to solving the "black box" problem. *Lex russica*, *73*(6), 139–148. (In Russ.). https://doi.org/10.17803/1729-5920.2020.163.6.139-148

Lang, H., & Shan, C. (2000). Bias phenomenon and compensation in multiple target tracking algorithms. *Mathematical and Computer Modelling*, *31*(8–9), 147–165. https://doi.org/10.1016/S0895-7177(00)00063-7

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review*, *48*, 105790. https://doi.org/10.1016/j.clsr.2023.105790

Lessig, L. (1999). *Code and other laws of cyberspace*. New York: Basic Books.

Li, Z. (2022). Affinity-based algorithmic pricing: A dilemma for EU data protection law. *Computer Law & Security Review*, *46*, 105705. https://doi.org/10.1016/j.clsr.2022.105705

Malyshkin, A. V. (2019). Integration of artificial intelligence into public life: some ethical and legal problems. *Vestnik of Saint Petersburg University. Law*, *10*(3), 444–460. (In Russ.). https://doi.org/10.21638/spbu14.2019.303

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, *38*(1), 101550. https://doi.org/10.1016/j.giq.2020.101550

Mourby, M., Ó Cathaoir, K., & Bjerre Collin, C. (2021). Transparency of machine-learning in healthcare: The GDPR & European health law. *Computer Law & Security Review*, *43*, 105611. https://doi.org/10.1016/j.clsr.2021.105611

Ostroumov, N. N. (2015). Legal regime of international air transportation. Moscow: Statut. (In Russ.).

Qiaochu, W., Yan, H., Stefanus, J., & Param Vir, S. (2020, July 15). *Algorithmic Transparency with Strategic Users*. http://dx.doi.org/10.2139/ssrn.3652656

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, *63*, 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Stahl, B. C., Rodrigues, R., Santiago, N., & Macnish, K. (2022). A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values. *Computer Law & Security Review*, *45*, 105661. https://doi.org/10.1016/j.clsr.2022.105661

Stefik, M. (1996). Letting loose the light: Igniting commerce in electronic publication. In M. Stefik (Ed.), *Internet dreams: Archetypes, myths, and metaphors* (pp. 219–253). Cambridge, MA: MIT Press.

Stöger, K., Schneeberger, D., Kieseberg, P., & Holzinger, A. (2021). Legal aspects of data cleansing in medical AI. *Computer Law & Security Review*, *42*, 105587. https://doi.org/10.1016/j.clsr.2021.105587

Su, Zh., Bentley, B. L., McDonnell, D., Cheshmehzangi, A., Ahmad, J., Šegalo, S., Pereira da Veiga, C., & Xiang, Yu-Tao. (2023). China's algorithmic regulations: Public-facing communication is needed. *Health Policy and Technology*, *12*(1), 100719. https://doi.org/10.1016/j.hlpt.2022.100719

Talapina, E. V. (2020). Algorithms and artificial intelligence in the human rights context. *Journal of Russian Law*, *10*, 25–39. (In Russ.). https://doi.org/10.12737/jrl.2020.118.

Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, *3*(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Wang, H. (2022). Transparency as Manipulation? Uncovering the Disciplinary Power of Algorithmic Transparency. *Philosophy & Technology*, *35*, 69. https://doi.org/10.1007/s13347-022-00564-w

Willems, J., Schmidthuber, L., Vogel, D., Ebinger, F., & Vanderelst, D. (2022). Ethics of robotized public services: The role of robot design and its actions. *Government Information Quarterly*, *39*(2), 101683. https://doi.org/10.1016/j.giq.2022.101683

Xu, J., Xiao, Yu., Wang, W. Hu., Ning, Yu., Shenkman, E. A., Bian, J., & Wang, F. (2022). Algorithmic fairness in computational medicine. *eBioMedicine*, *84*, 104250. https://doi.org/10.1016/j.ebiom.2022.104250

Zharova, A. (2019). Ensuring the information security of information communication technology users in Russia. *International Journal of Cyber Criminology*, *13*(2), 255–269. DOI: https://doi.org/10.5281/zenodo.3698141

Zharova, A. (2020). The protect mobile user data in Russia. *International Journal of Electrical and Computer Engineering*, *10*(3), 3184–3192. https://doi.org/10.11591/ijece.v10i3.pp3184-3192

Zhu, H., Sallnäs Pysander, E.-L., & Söderberg, I.-L. (2023). Not transparent and incomprehensible: A qualitative user study of an AI-empowered financial advisory system. *Data and Information Management*, 100041. https://doi.org/10.1016/j.dim.2023.100041

Author information

Anna K. Zharova – Dr. Sci. (Law), Associate Professor, Senior Researcher, Institute of State and Law of the Russian Academy of Sciences
Address: 10 Znamenka Str., 420100 Moscow, Russian Federation
E-mail: anna_jarova@mail.ru
ORCID ID: https://orcid.org/0000-0002-2981-3369
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56964137900
WoS Researcher ID: https://www.webofscience.com/wos/author/record/H-4012-2015
Google Scholar ID: https://scholar.google.com/citations?user=g8ij3BsAAAAJ
RSCI Author ID: https://elibrary.ru/author_items.asp?authorid=151076

Conflict of interest

The author is an Editor-in-Chief of the Journal; the article has been reviewed on general terms.

Financial disclosure

The research was not sponsored.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – May 22, 2023
Date of approval – August 21, 2023
Date of acceptance – November 30, 2023
Date of online placement – December 15, 2023