ГОСУДАРСТВО И ПРАВО, 2023, № 7, с. 129–138
STATE AND LAW, 2023, No. 7, pp. 129–138

ИНФОРМАЦИОННОЕ ПРАВО И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

БОЛЬШИЕ ДАННЫЕ И ПРАВА ЧЕЛОВЕКА: НА ПУТИ К ПРАВОВОМУ РЕГУЛИРОВАНИЮ

© 2023 г. Э. В. Талапина

Институт государства и права Российской академии наук, г. Москва

E-mail: talapina@mail.ru

Поступила в редакцию 09.01.2023 г.

Аннотация. В отсутствие общепризнанного определения больших данных их признаки описаны в современной литературе. Обработка больших данных пока находится вне зоны действия законодательства, поэтому наиболее значимую юридическую проблему представляет «соприкосновение» с персональными данными, в отношении которых правила уже установлены. Любые нарушения в этой сфере напрямую затрагивают права человека, особенно право на частную жизнь. Применительно к большим данным конфиденциальность (частная жизнь) все чаще рассматривается как экономическое право, а персональные данные – как имеющие экономическую ценность. Необходимость соблюдения права на конфиденциальность ставит субъекта данных в центр правовой конструкции законного доступа к персональным данным, который должен выразить информированное согласие на операции с персональными данными.

В условиях оборота больших данных необходимо подобрать более подходящую и динамичную модель информированного согласия. Хотя данные как таковые вряд ли могут быть собственностью, это не означает, что они не должны защищаться с помощью других механизмов, например, направленных на обеспечение подконтрольности и подотчетности. Именно развитие механизмов подконтрольности становится одним из главных направлений в регулировании больших данных.

Ключевые слова: конфиденциальность, информированное согласие, персональные данные, большие данные, цифровизация, алгоритм, искусственный интеллект.

Цитирование: *Талапина Э. В.* Большие данные и права человека: на пути к правовому регулированию // Государство и право. 2023. № 7. С. 129–138.

DOI: 10.31857/S102694520026812-4

BIG DATA AND HUMAN RIGHTS: TOWARD LEGAL REGULATION

© 2023 E. V. Talapina

Institute of the State and Law of the Russian Academy of Science, Moscow

E-mail: talapina@mail.ru

Received 09.01.2023

Abstract. In the absence of a universally accepted definition of Big Data, its attributes are described in the current literature. The Big Data processing is still outside the legal frameworks, so the most significant legal problem is the "contact" with personal data, for which rules have already been established. Any violations in this area directly affect Human Rights, especially the right to privacy. In the case of Big Data, privacy is increasingly seen as an economic right and personal data as having economic value. The need to respect the right to privacy puts the data subject at the center of the legal construction of legal access to personal data who expresses informed consent to operations with his personal data.
In the context of Big Data processing, it is necessary to select a more appropriate and dynamic model of informed consent. Although data as such is unlikely to be a right of property object, it should be protected through other mechanisms, such as those aimed at ensuring accountability. It is the development of accountability mechanisms that is becoming one of the main directions in the Big Data regulation.

Key words: privacy, informed consent, personal data, Big Data, digitalization, algorithm, artificial intelligence.

For citation: *Talapina, E.V. (2023).* Big Data and human rights: toward legal regulation // Gosudarstvo i pravo=State and Law, No. 7, pp. 129–138.

Природа больших данных

Хотя точного определения больших данных не существует, их признаки хорошо описаны в современной литературе. Они заключаются в значительном увеличении объема данных, которые можно генерировать и хранить, в скорости, с которой данные могут быть предоставлены для принятия решений в режиме реального времени, в разнообразии форматов, в которых данные могут быть приняты. Исходя из этого, изначально преобладал подход, определяющий три ключевые характеристики больших данных (получивший наименование три "V"): большой объем (Volume), разнообразие данных (Variety), высокая скорость их изменения (Velocity). Затем выделили 4 "V" (Value – ценность данных) и 5 "V" (Veracity – достоверность данных): достоверность данных, связанная с определенными типами данных, а также возможность извлечения ценности путем определения того, что является ценным, а затем преобразования и извлечения для анализа. Соответственно, термин «большие данные» стал определяться через пять атрибутов, а именно: объем, скорость, разнообразие, достоверность и ценность.

Но удивительным образом дальнейшая практика и изучение больших данных открывают все новые характеристики, число которых теперь доводится до семи и даже восьми (Validation – проверка данных)[1], изменчивость (Variability) и визуализация (Visualization)[2], возможно, это еще не предел. Такая динамика свидетельствует о громадном неизученном потенциале технологий, которые кумулируют как потенциальные выгоды для человечества, так и большие риски.

Большие данные далеко не однородны. Они собираются из разных источников, таких как открытые данные, социальные сети, интернет вещей, персональные данные, коммерческие транзакции и пр.[3] Разнообразие больших данных не может не обострить потребность в их классификации. Причем в данном случае смысл такой классификации далеко не только теоретический, но и сугубо практический, поскольку позволяет выстроить в будущем системное регулирование в области использования больших данных. Интересную попытку систематизации содержит

[1] См.: *Bagnoli V.* Competition for the Effectiveness of Big Data Benefits // IIC – International Review of Intellectual Property and Competition Law. 2015. Vol. 46. P. 629–631; *Fredriksson C., Mubarak F., Tuohimaa M., Zhan M.* Big Data in the Public Sector: A Systematic Literature Review // Scandinavian Journal of Public Administration. 2017. Vol. 21(3). P. 39–61.

[2] См.: *Rijmenam M.* Why the 3V's Are Not Sufficient To Describe Big Data // Datafloq. URL: https://datafloq.com/read/3vs-sufficient-describe-big-data/166 (дата обращения: 04.01.2023).

[3] См.: *Шаталова В.В., Лихачевский Д.В., Казак Т.В.* Большие данные: как технологии BIG DATA меняют нашу жизнь // Big Data and Advanced Analytics. 2021. № 7-1. С. 189.

исследовательский отчет Управления Великобритании по конкуренции и рынкам о коммерческом использовании потребительских данных (2015)[4]. Несмотря на то что он касается отдельных секторов экономики (автострахование, розничная торговля и игровые приложения), в отчете была предпринята попытка дать представление о потребительских данных в качестве персональных и неперсональных данных, псевдонимных и агрегированных. В нем также рассматривались способы, с помощью которых собираются потребительские данные, – выводятся ли они, заявляются явным образом или собираются путем наблюдения за взаимодействиями пользователей.

Этот посыл весьма важен, поскольку показывает, насколько много граней у персональных данных и как легко эти грани преступить. Поведенческие данные, данные об использовании и содержании, данные об опыте, технические данные и данные о местоположении – все это подкатегории персональных данных, хотя и косвенные, по сравнению с более прямыми, или очень чувствительными, персональными данными. Агрегированные данные относятся к категории данных, вытекающих из любой из вышеперечисленных категорий[5].

С юридической точки зрения необходимо отметить, что из всего перечисленного многообразия лишь персональные данные регулируются законодательством. Большие данные, которые не содержат непосредственно идентифицируемых персональных данных или данных, которые могут быть объединены с другими данными для идентификации личности, не подпадают под законодательство о защите данных. В итоге создается ситуация неурегулированности использования больших данных. Поэтому в принципе большие данные могут бесконтрольно использоваться в деятельности коммерческих и государственных организаций, что и происходит. Помимо сомнительной законности обогащения риски коммерческого использования персональных данных включают в себя профилирование и ограниченный доступ к продуктам и услугам в связи с соответствующим профилированием. В качестве некоего ограничителя могут выступать этические требования, существующие в некоторых отраслях, но это, во-первых, не обязывает юридически, а во-вторых, распространяется на весьма ограниченные сферы.

Таким образом, наиболее заметные юридические проблемы лежат в плоскости «соприкосновения» персональных данных, обращение с которыми уже длительное время хорошо отрегулировано, и больших данных, оборот которых находится вне зоны действия законодательства.

Кроме того, зачастую сложность представляет сама ***квалификация данных в качестве персональных***. Иногда неличные, казалось бы, данные (переменные данные, не являющиеся персональными, – данные о местоположении (геолокация) или метаданные, основанные на использовании телефонных услуг) могут быть использованы для извлечения персональных данных.

В некоторых случаях квалификация данных в качестве персональных зависит от цели использования, которую непросто доказать. Например, европейское Руководство по обработке персональных данных посредством видеоаппаратуры от 29 января 2020 г.[6] устанавливает, что лицо, использующее персональное устройство или экшн-камеру, прикрепленную к спортивному снаряжению, для записи действий во время отпуска, может быть защищено в рамках исключений из Общего регламента по защите данных Европейского Союза (далее – GDPR), касающихся домашних хозяйств (ст. 2(2)(c)), даже если на заднем плане ведется запись третьих лиц. Однако существует особый акцент на том, что исключение применяется только к случаям, когда запись показывается друзьям и членам семьи. Руководство ссылается на решение суда Европейского Союза 2003 г., в котором говорится, что загрузка видео в Интернет и предоставление доступа к данным неопределенному кругу лиц не подпадает под действие освобождения от уплаты налогов на домашнее хозяйство.

В то же время симметрично жесткие рамки установлены и для видеонаблюдения в целях безопасности. Само по себе «видеонаблюдение в целях безопасности» не будет рассматриваться как достаточно конкретная цель.

Эти нюансы, проработанные в европейском регулировании, могут быть полезны для России. Известно, насколько «неаккуратно» в нашей стране используются видеоустройства, с многочисленными нарушениями прав третьих лиц на их персональные данные (записи с видеорегистраторов в автомобилях, видеосъемка в общественных местах, видео с камер наблюдения практически беспрепятственно попадают в Сеть и распространяются). Самое печальное, что основная масса людей

[4] См.: DotEcon & Analysys Mason (2015), The Commercial Use of Consumer Data A research report for the CMA. URL: https://www.gov.uk/cma-cases/commercial-use-of-consumer-data (дата обращения: 04.01.2023).

[5] См.: *Chirita A.D.* The Rise of Big Data and the Loss of Privacy // Personal Data in Competition, Consumer Protection and Intellectual Property Law / ed. by M. Bakhoum et al. Springer, 2018. P. 175.

[6] См.: Guidelines 3/2019 on processing of personal data through video devices, 29 January 2020. URL: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201903_video_devices.pdf (дата обращения: 04.01.2023).

в таких ситуациях даже не подозревает о том, что при этом нарушаются их права.

Право на частную жизнь и персональные данные

Одним из главных «страдающих» прав человека в условиях применения цифровых технологий является право на частную жизнь. Существует два способа, которыми раскрытие частной информации может нанести вред субъекту этой информации. Первый – это уменьшение свободы: оно способно привести к вмешательству других людей или к формированию у них желания вмешаться. Вмешательство в частную жизнь может не создавать полных препятствий, а уменьшать свободу действий (фокус на внешних препятствиях для будущих действий). Второй способ, которым раскрытие частной информации может нанести вред субъекту, заключается в уменьшении известной свободы: убеждения человека о его свободе и несвободе ухудшаются[7]. Таким образом, вторжение в частную жизнь может уменьшить свободу человека или же его знания о свободе, и оба эти фактора наносят ущерб интересам свободы: раскрытие частной информации приводит к изменениям в негативной свободе человека, а также к изменениям в знаниях, которые человек имеет относительно степени своей негативной свободы[8].

Цифровые технологии несколько поколебали эту классическую конструкцию. С одной стороны, вмешательств в частную жизнь и утечек персональных данных стало так много, что позволяет некоторым утверждать об исчезновении приватности как социальной нормы. С другой стороны, рост подобных утечек не позволяет оставить людей вовсе без защиты, и фокус науки и практики смещается на возмещение причиненного ущерба.

Вопросы расчета и возмещения ущерба справедливо адресуются судам. Как констатируют исследователи, суды придерживаются мнения, что риск будущей кражи личных данных не представляет собой «фактический» ущерб[9]. При этом зависимость судов от принципа «не навреди» в недавних делах о нарушении данных[10] должна быть пересмотрена, иначе дело не сдвинется с мертвой точки. Кроме того, формируется опасная тенденция – вместо того чтобы присуждать убытки в делах, связанных с утечкой данных, становится достаточно лишь проверить соблюдение требования об уведомлении клиентов об утечке данных.

На этом фоне российские законодательные инициативы выглядят несколько старомодно, преследуя цель стимулирования бизнеса посредством ужесточения ответственности за утечки, – установление крупных штрафов призвано повысить уровень обеспечения безопасности обработки персональных данных, как считает Минцифры, разработавшее соответствующий законопроект[11].

В основе рассуждений об ответственности за утрату данных лежит признание ***экономической ценности персональных данных***. Действительно, применительно к большим данным конфиденциальность (частная жизнь) все чаще рассматривается как основополагающее экономическое право[12]. А неприкосновенность частной жизни как экономическое право нуждается в нормативном обрамлении, причем не только в законодательстве о защите данных, но и в законодательстве о защите потребителей и др.

На практике частная жизнь действует как щит, который ограничивает сбор или использование персональных данных о человеке. Исключительные требования ст. 8 Хартии ЕС, а именно справедливость и законность обработки данных для определенной цели, прозрачность (включая право на доступ к данным) следует рассматривать как конституционное измерение неприкосновенности частной жизни. В европейском законодательстве введен ряд прав, гарантирующих сохранение определенного контроля субъектов данных над своими персональными данными. Так, европейский GDPR предусматривает права на доступ, исправление, стирание и переносимость данных. Многие аналогичные нормы есть в России. Все это ставит субъекта данных в центр правовой конструкции законного доступа к персональным данным, который выражает информированное согласие на операции с персональными данными.

Информированное согласие субъектов данных

Предусмотренное законодательством о персональных данных информированное согласие имеет исторические корни. Например, информированное согласие существует в медицине и означает практику информирования пациентов об их здоровье и медицинских условиях, а также предоставление им возможности выбора между различными терапевтическими альтернативами. Изначально в корпусе Гиппократа не было обязательства говорить правду: пациент не нуждается в информировании. Более того, часто информация должна быть скрыта, чтобы не напугать пациента. И даже нет никакого упоминания о возможности выбора у пациента: врач является обладателем медицинских

[7] См.: *De Bruin B.* The liberal value of privacy // Law and Philosophy. 2010. Vol. 29. P. 532.

[8] См.: ibid. P. 511.

[9] См.: ibid. P. 532.

[10] См.: ibid. P. 533.

[11] См.: URL: https://www.kommersant.ru/doc/5747151 (дата обращения: 04.01.2023).

[12] См.: *Chirita A.D.* Op. cit. P. 159.

знаний и поэтому знает, что лучше для пациента. Однако со временем эта конструкция пошатнулась; первые критики традиционной концепции Гиппократа появились в середине XIX в. в США, а самые ранние судебные иски были поданы во втором десятилетии XX в.[13] До Европы эта тенденция дошла гораздо позднее – к примеру в Италии внедрение информированного согласия было медленным и постепенным (90-е годы XX в.) и названо «тихой революцией».

Для темы нашего исследования главный вывод из истории информированного согласия в том, что для его дачи пациент должен быть «компетентным». С одной стороны, это требует создания целой системы доступного информирования по специфическим вопросам, в которых человек не является специалистом. С другой стороны, это исключает для определенных категорий лиц возможность дать действительное информированное согласие (например, младенцы или люди с деменцией). И здесь снова с моральной точки зрения понятие «человек» становится основополагающим для полноценного понимания информированного согласия. Информированное согласие (как и заблаговременное распоряжение о нем) позволяет человеку выразить свою индивидуальность и, следовательно, свое равенство в глазах закона.

Применительно к персональным данным получение согласия субъектов данных обеспечивает сохранение их самостоятельности и возможности вмешиваться в процесс принятия решений относительно использования их данных. Но, как неоднократно отмечалось, и нами в том числе, конструкция согласия рассчитана на линейные отношения, когда субъект данных сам предоставляет свои данные обработчику, обозначая цель (например, в административных отношениях, заполняя заявления и прочие формы). Использование больших данных, в которых могут содержаться неанонимизированные персональные данные, и даже персональные данные, выведенные путем сопоставления ряда других (агрегированные), принципиально меняет условия игры. О таких использованиях человек может попросту не знать, не догадываться, не прогнозировать и пр.

Получается, что, учитывая разнообразные источники больших данных и тот факт, что будущая полезность больших данных обычно не определена на момент получения информированного согласия, согласие не может быть действительно информированным из-за трудности прогнозирования и информирования субъектов данных о будущем использовании, равно как о последствиях использования данных в будущем[14]. Таким образом, необходимо подобрать более подходящую и ***динамичную модель информированного согласия***.

Неплохой пример можно найти в литературе применительно к использованию больших данных в научных медицинских целях, где описаны три модели согласия – широкое согласие, согласие с отказом и динамическое согласие.

При широком согласии субъекты данных акцептуют ряд возможных исследований, которые могут быть проведены с использованием их данных в отношении конкретной области или направления исследований. Для того чтобы широкое согласие было действительным, в качестве гарантии должны существовать соответствующие комитеты по проверке, которые обеспечат защиту прав субъектов данных. Модель отказа от использования предполагает, что данные могут быть использованы, если субъект данных явно от этого не отказался. Проблема этой модели заключается в том, что субъекты данных могут быть недостаточно информированы об условиях использования, особенно в коммерческих базах данных или социальных сетях. Динамическая модель позволяет субъектам данных обновлять свое согласие на постоянной основе. Хотя эта модель в основном применяется в биобанкинге, она также может быть использована в обстоятельствах, которые влекут за собой многократное и разнообразное использование данных (большие данные), где могут потребоваться различные виды согласия с течением времени. Последняя модель обеспечивает открытую коммуникацию с субъектами данных.

Право «собственности» на данные

В XX в. человек и его персональные данные, использовавшиеся в основном для административных целей, не вызывали вопросов о собственности. Скорее наоборот, данные, поступившие в органы государственной власти, рассматривались в качестве некоего общественного ресурса (зачастую неофициально). Хотя в России, к примеру, в первом Законе об информации признавалось право собственности на информационные ресурсы, причем во всех формах[15], но эти нормы было трудно применять на практике. Другой показательный пример неудачного законодательного подхода – это проблемы, с которыми столкнулась Исландия в 1998 г. Там в связи с объявлением медицинских карт, включающих медицинские, генетические и генеалогические данные, в качестве

[13] См.: Schloendorf v. Society of New York Hospital. 1914. 105 N.E. 92, NY; Salgo v. Leland Stanford, Jr. University Board of Trustees. 1957. 154 Cal. App. 2d 560 317 P. 2d 170.

[14] См.: *Mittelstadt D.B., Floridi L.* The ethics of big data: current and foreseeable issues in biomedical contexts // Sci Eng Ethics. 2016. Vol. 22. P. 312.

[15] См.: Федеральный закон от 20.02.1995 г. № 24-ФЗ «Об информации, информатизации и защите информации» // СЗ РФ. 1995. № 8, ст. 609. Утратил силу.

национального ресурса, который принадлежит исландскому правительству и может быть доступен частной промышленности без согласия отдельных лиц, возникла национальная и международная оппозиция против такого неподобающего способа, которым правительство Исландии решало вопрос о праве собственности на данные. В результате проект потерпел крах в 2003 г.

Как видим, несмотря на неоднократные попытки, право собственности в информационной сфере с трудом поддается традиционному урегулированию. Экономический контекст, а именно возможность использования данных коммерческими организациями для извлечения прибыли, только обостряет проблему. С распространением технологий больших данных вопросы собственности на данные кажутся еще более запутанными. С одной стороны, очевидна экономическая стоимость персональных данных, с другой – до сих пор непонятно, каким образом направить отношения их обладателя с заинтересованными лицами в юридическое русло. Отправной точкой для поиска можно считать замечание ЮНЕСКО о том, что концепция собственности больше не является адекватной нормативной базой в эпоху больших данных[16]. Все большую популярность набирает мнение, что право собственности – это концепция, которая плохо подходит для регулирования конкурирующих прав в больших данных[17]. Так, авторское право распространяется только на обработанные данные, а необработанные данные не являются оригинальным творением[18].

Немаловажное значение для ответа на поставленный вопрос имеет и преследуемая цель – собираемся ли мы ускорить развитие экономики данных за счет их повышенной конвертации или уделить внимание защите прав обладателей данных? Хотя эти цели несколько разнонаправлены, хотелось бы найти усредненный путь: стремление использовать весь потенциал больших данных должно в то же время сопровождаться уважением прав других пользователей на доступ к информации.

Одним из оригинальных предложений представляется отход от риторики собственности вообще. С учетом того, что права собственности на данные еще больше затруднят обмен данными (а это приведет к тому, что все заинтересованные стороны будут предъявлять претензии на данные), в науке предлагается кардинально сменить парадигму от владения к опеке (Paradigm Shift from Ownership to Custodianship). Соответствующая опека над большими данными необходима для обеспечения того, чтобы субъекты данных сохраняли определенный контроль над доступом и будущим использованием своих данных, при этом делегируя принятие решений по некоторым вопросам хранителям данных. Такое делегированное принятие решений порождает права опеки, а не права собственности на данные[19].

Подобное предложение сформулировано применительно к использованию персональных данных в целях медицинских исследований, поскольку ни правовая концепция собственности, ни даже использование договорных условий не решили проблем в исследованиях в области здравоохранения с использованием больших данных. Автор мотивирует свое предложение тем, что опека гораздо шире, чем юридическая концепция собственности, и используется для обеспечения того, чтобы все заинтересованные стороны признавали и выполняли свои этические обязательства служить наилучшим интересам биомедицинских исследований.

Система этических рамок опекунства подразумевает признание данных в качестве дара от субъектов данных, которые будут использоваться с их согласия для развития науки на благо общества и не являться собственностью исследователей, принимающих институтов или финансирующих организаций. Исследователи, получившие данные, считаются хранителями этих данных. Обязанности хранителя подразумевают соблюдение строгих этических и нормативных требований, таких как предоставление точных и своевременных данных и обеспечение неприкосновенности частной жизни и конфиденциальности субъектов данных.

Это – аргумент в пользу вывода о том, что ***данные как таковые вряд ли могут быть собственностью***. Однако это не означает, что они не должны защищаться с помощью других механизмов, например, направленных на обеспечение подотчетности, вместо предоставления прав собственности. Именно развитие механизмов подконтрольности, подотчетности становится одним из главных направлений в регулировании больших данных.

Подконтрольность как способ соблюдения прав человека при обработке больших данных

Проблема обеспечения подконтрольности в использовании технологий обретает особое значение. В условиях обрывочного правового регулирования (а в отношении больших данных, напомним, его практически нет) объяснение сути происходящего – почти единственно верный путь, позволяющий защищать права человека. Среди многих

[16] См.: UNESCO (2017) Report of the International Bioethics Committee of UNESCO (IBC) on big data and health. Paris, UNESCO. URL: https://unesdoc.unesco.org/ark:/48223/pf0000248724 (дата обращения: 04.01.2023).

[17] См.: *Andanda P.* Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research // IIC. 2019. Vol. 50. P. 1072.

[18] См.: *Froehlich A., Tăiatu C.M.* Space in Support of Human Rights. Cham, 2020. P. 36.

[19] См.: *Andanda P.* Op. cit. P. 1075.

недостатков и рисков больших данных (нарушение конфиденциальности, предвзятость и дискриминация, ошибочность данных и пр.) масштабное воздействие на поведение людей становится ключевым, поскольку целится в самое главное – автономию человека.

Как справедливо отмечают исследователи, первоначально наблюдался этап энтузиазма в отношении больших данных, которые предполагались объективными и достоверными[20]. Однако со временем концепция объективности больших данных подверглась критике. Действительно, осмысление данных всегда предполагает человеческие предубеждения, поскольку осуществляется с использованием определенной (выбранной человеком) оптики, которая влияет на интерпретацию данных[21]. О необъективности алгоритмов, которые отражают ценности разработчиков и в результате могут быть дискриминационными, уже немало написано[22]. В Европейском Союзе GDPR частично решает эту проблему, так как физические лица имеют право возражать против автоматизированных решений и требовать человеческого вмешательства для пересмотра решения. Аналогичная норма существует в российском законодательстве.

Говоря о воздействии алгоритмов на поведение людей, необходимо изучить и понять любую роль, которую алгоритмы и технологические платформы играют в усилении социальной фрагментации, – например, подавая исключительно консервативные новости для консервативных пользователей и либеральные новости для либеральных пользователей. Весьма показателен здесь американский опыт. В 2019 г. в США был разработан закон о прозрачности пузырей фильтров (Filter Bubble Transparency Act), специально для потребителей социальных сетей и политических новостей (отдельно от закона о подотчетности алгоритмов). Данный закон имеет заявленную цель гарантировать американцам право на взаимодействие с (медиа) платформой, не подвергаясь манипулированию алгоритмами, управляемыми данными конкретного пользователя.

В своих попытках законодательно закрепить использование персональных данных интернет-платформами авторы законопроекта проводят различие между двумя типами «пользовательских» данных: те, которые были «явно предоставлены пользователем платформе» для целей алгоритмической системы ранжирования, и те, которые не были предоставлены. Законопроект уточняет, что платформам для определения контента разрешено использовать список аккаунтов, на которые человек подписан в социальных сетях. Однако любая фильтрация, упорядочивание или ранжирование контента (за исключением хронологического упорядочивания) требуют, чтобы платформы показывали соответствующий значок рядом с лентой контента. Этот значок будет служить двум основным целям: 1) информировать пользователей о том, что их лента фильтруется на основе поведенческих данных конкретного пользователя и 2) позволит пользователям выбирать между алгоритмически ранжированной лентой и хронологической лентой.

Требование прозрачности, предложенное в законопроекте FBTA, подразумевает, что поведение людей, использующих социальные медиа и контент-платформы, будет иным, если они будут обладать большей информацией. Информирование пользователей о том, что их ленты подвергаются алгоритмической фильтрации, могло бы заставить их более внимательно относиться к тому, на какую информацию они «кликают», что им нравится, на что они реагируют, либо сделать выбор в пользу другого (возможно, чисто хронологического) типа фильтрации контента.

Оценивая законопроект, исследователи отмечают, что поведенческие вмешательства, специально разработанные для смягчения эффекта «пузыря фильтров», имеют ограниченное действие, а компонент прозрачности в предложении FBTA будет иметь незначительный эффект в отношении изменения поведения потребителей на платформах цифрового контента[23].

Наблюдение об очевидной неоднородности алгоритмических эффектов говорит о том, что обсуждение пузырей цифровых фильтров без систематических и контекстуальных нюансов может привести нас к упрощенным выводам. На самом деле между факторами в алгоритмической социальной системе могут существовать сложные взаимодействия. Это было показано путем симулятивного анализа. Как и в случае с алгоритмами рекомендаций в реальной жизни, на рекомендации, которые получает пользователь в этой симуляции, влияют также рекомендации других пользователей платформы. Это привносит сложный набор динамики, который затрудняет априорное предсказание того, как один алгоритм повлияет на результаты работы системы по сравнению с другим.

[20] См.: Интерпретация и применение больших данных в юриспруденции и юридической практике / науч. ред. Ю.А. Тихомиров. М., 2021. С. 20.

[21] См.: там же. С. 31.

[22] См., напр.: *Талапина Э.В.* Обработка данных при помощи искусственного интеллекта и риски дискриминации // Право. Журнал ВШЭ. 2022. № 1. С. 4–27.

[23] См.: *Hosanagar K., Miller A.P.* Who Do We Blame for the Filter Bubble? On the Roles of Math, Data, and People in Algorithmic Social Systems // After the Digital Tornado: networks, algorithms, humanity / ed. by K. Werbach. Cambridge University Press, 2020. P. 111.

В реальном мире алгоритмы еще более непрозрачны, данные более массивны, а пользователи демонстрируют более сложное поведение. Все эти факторы только увеличивают сложность социальной системы и указывают на необходимость более глубокого понимания и учета тонкостей, связанных с взаимодействием всех вовлеченных факторов. К примеру, опытным путем было установлено, что один и тот же алгоритм может иметь кардинально разные эффекты в зависимости от контекста, в котором он применяется, а одни и те же входные данные могут иметь различные результаты в зависимости от алгоритма, который действует на эти данные. Таким образом, если мы хотим успешно достигать при помощи алгоритмических систем социально полезных результатов, важно оценить сложность этих систем и не допустить упрощенных обобщений относительно происходящей в них динамики[24].

Во многих научных исследованиях применительно к владельцу больших данных используется термин «власть». Владение и контроль над большими данными, а также решение о том, кто еще их получит, автоматически означает власть для тех, кто контролирует данные[25]. Естественно, что любая власть нуждается в сдерживании и ограничении. Пока же в отношении больших данных действуют лишь этические ограничители, довольно бессильные перед лицом коммерческих притязаний и жажды прибыли. Кроме того, даже в случае запуска механизма ответственности за этическое управление большими данными субъекта такой ответственности трудно определить из-за сложности и интеграции последних.

Вместо обобщения: некоторые подходы к регулированию больших данных

Цифровой мир, в котором используются данные, создает угрозу того, что субъекты данных могут потерять контроль над своими данными. На настоящий момент использование больших данных не регулируется; как частные субъекты, так и государство могут использовать их для принятия решений (в отношении о своей деятельности или людей в самых разных статусах – клиентов, граждан, работников). По сути, единственным ограничителем подобной практики служат этические нормы, не обязательные к применению. Ситуация неурегулированности сама по себе провоцирует риски для прав и свобод человека. При этом наибольшую угрозу неприкосновенности частной жизни и защите данных представляет «связка» между технологиями искусственного интеллекта и больших данных, поскольку обе эти технологии используются для профилирования граждан. Кроме того, незнание граждан о том, что их персональные данные обрабатываются в составе больших, фактически блокирует любую защиту прав и свобод, которые могут быть при этом нарушены.

Как представляется, есть как минимум три проблемы, связанные с использованием больших данных. К ним относятся: доступ к данным, согласие субъектов данных на обработку данных, а также претензии на право «собственности», которые могут препятствовать обмену данными. Это тесно связано с контролем субъектов данных над тем, кто может получать доступ, использовать и делиться информацией. Вокруг решения перечисленных проблем и должно выстраиваться будущее правовое регулирование.

Для организации регулирования необходимо прежде всего определить базовые понятия. Юридическая наука и тем более практика нуждаются в четкости определений. Только так можно рассчитывать на справедливое правоприменение и судопроизводство. Выход видится в детальной дифференциации и категоризации данных. Применительно к большим данным особенно актуально определить правовые режимы агрегированных данных, а также персональных данных в составе больших данных.

Эпоха больших данных затрудняет субъектам данных возможность предвидеть конкретные будущие виды использования, в основном из-за сложной взаимосвязи между многочисленными и меняющимися источниками данных. Поэтому, хотя согласие может в некоторой степени решить проблемы, связанные с характером использования больших данных, оно не решает в достаточной степени проблему вторичного использования данных. Стоит также выйти за рамки согласия и перейти к более широкой системе подконтрольности, которая учитывает такие вопросы, как вред и оценка риска.

Как мы установили, законодательство о защите данных сосредоточено на персональных данных, но большие данные обычно не включают в себя непосредственно идентифицирующую личность информацию. Соответственно, на данный момент большие данные не попадают под действие законодательства. На них не распространяется защита данных, так как объектом являются группы, а не идентифицируемый индивид. Это представляет собой значимую юридическую проблему, поскольку актуальное законодательство не обеспечивает должный нормативный надзор, необходимый для использования больших данных.

Сегодня понятие «этика алгоритмов» (которые обычно используют большие данные для формирования предположений о людях) уже достаточно устоялось и реализуется на практике. Однако уровня этического сдерживания явно не хватает.

[24] См.: *Hosanagar K., Miller A.P.* Op. cit. P. 117, 118.

[25] См.: *Roche J., Jamal A.* A Systematic Literature Review of the Role of Ethics in Big Data // Cybersecurity, Privacy and Freedom Protection in the Connected World. Proceedings of the 13th International Conference on Global Security, Safety and Sustainability. London, January 2021 / ed. by H. Jahankhani, A. Jamal, S. Lawson. Cham, 2021. P. 337.

Одним из выходов из ситуации неурегулированности больших данных может стать применение т.н. принципа субсидиарности – путем дополнения этики больших данных системой управления данными, когда подотчетные должностные лица наделяются ответственностью за конкретные наборы данных, чтобы обеспечить их надлежащее управление. Эти обязанности могут быть расширены за счет добавления этических соображений в отношении сбора и обработки набора данных. В литературе рекомендуется, чтобы разработчики политики и руководители высшего звена приняли этот подход в качестве временной меры до того, как будет разработано более общее законодательство[26]. Другой вариант субсидиарности, на наш взгляд, – дополнить этику больших данных полноценным информационным самоопределением граждан; здесь центр тяжести будет не столько на ответственном обработчике персональных данных, сколько (упреждающе) на самом их обладателе. Возможен и третий вариант («тройственный союз») – распределение ответственности между обработчиком персональных данных, их обладателем и обработчиком больших данных, что потребует более филигранной нормативной детализации и точной реализации на практике.

Каким образом все это может быть оформлено законодательством? Пока ясно то, что оно должно учитывать нюансы разных уровней, поскольку изначально комплексно. Для иллюстрации сложности проблемы Буртшер и Фритц сравнили большие данные с глыбой мрамора, над которой работают несколько резчиков, так что различные правовые концепции, включая конфиденциальность данных, права на базы данных, права интеллектуальной собственности, антимонопольное законодательство, а также основные гражданские права собственности и владения играют определенную роль при работе с юридически чуждыми большими данными, при этом каждая из них затрагивает лишь отдельные части этих данных[27]. Создать комплексное законодательство – перспективная задача для юристов.

СПИСОК ЛИТЕРАТУРЫ

1. Интерпретация и применение больших данных в юриспруденции и юридической практике / науч. ред. Ю.А. Тихомиров. М., 2021. С. 20, 31.
2. *Талапина Э.В.* Обработка данных при помощи искусственного интеллекта и риски дискриминации // Право. Журнал ВШЭ. 2022. № 1. С. 4–27.
3. *Шаталова В.В., Лихачевский Д.В., Казак Т.В.* Большие данные: как технологии BIG DATA меняют нашу жизнь // Big Data and Advanced Analytics. 2021. № 7-1. С. 189.
4. *Andanda P.* Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research // IIC. 2019. Vol. 50. P. 1072, 1075.
5. *Bagnoli V.* Competition for the Effectiveness of Big Data Benefits // IIC – International Review of Intellectual Property and Competition Law. 2015. Vol. 46. P. 629–631.
6. *Burtscher B., Fritz G.* (2015) Big data: who owns and who may use and exploit big data? URL: https://www. lexology.com/library/detail.aspx?g=77ab3ffb-8f25-469c-ad80-d14bbcee9551 (дата обращения: 04.01.2023).
7. *Chirita A.D.* The Rise of Big Data and the Loss of Privacy // Personal Data in Competition, Consumer Protection and Intellectual Property Law / ed. by M. Bakhoum et al. Springer, 2018. P. 159, 175.
8. *De Bruin B.* The liberal value of privacy // Law and Philosophy. 2010. Vol. 29. P. 511, 532, 533.
9. DotEcon & Analysys Mason (2015), The Commercial Use of Consumer Data A research report for the CMA. URL: https://www.gov.uk/cma-cases/commercial-use-of-consumer-data (дата обращения: 04.01.2023).
10. *Fredriksson C., Mubarak F., Tuohimaa M., Zhan M.* Big Data in the Public Sector: A Systematic Literature Review // Scandinavian Journal of Public Administration. 2017. Vol. 21(3). P. 39–61.
11. *Froehlich A., Tăiatu C.M.* Space in Support of Human Rights. Cham, 2020. P. 36.
12. *Hosanagar K., Miller A.P.* Who Do We Blame for the Filter Bubble? On the Roles of Math, Data, and People in Algorithmic Social Systems // After the Digital Tornado: networks, algorithms, humanity / ed. by K. Werbach. Cambridge University Press, 2020. P. 111, 117, 118.
13. *Mittelstadt D.B., Floridi L.* The ethics of big data: current and foreseeable issues in biomedical contexts // Sci Eng Ethics. 2016. Vol. 22. P. 312.
14. *Rijmenam M.* Why the 3V's Are Not Sufficient To Describe Big Data // Datafloq. URL: https://datafloq.com/read/3vs-sufficient-describe-big-data/166 (дата обращения: 04.01.2023).
15. *Roche J., Jamal A.* A Systematic Literature Review of the Role of Ethics in Big Data // Cybersecurity, Privacy and Freedom Protection in the Connected World. Proceedings of the 13th International Conference on Global Security, Safety and Sustainability. London, January 2021 / ed. by H. Jahankhani, A. Jamal, S. Lawson. Cham, 2021. P. 337, 340.

REFERENCES

1. Interpretation and application of Big Data in jurisprudence and legal practice / scientific ed. Yu. A. Tikhomirov. M., 2021. P. 20, 31 (in Russ.).
2. *Talapina E.V.* Data processing using artificial intelligence and the risks of discrimination // Law. HSE Journal. 2022. No. 1. P. 4–27 (in Russ.).
3. *Shatalova V.V., Likhachevsky D.V., Kazak T.V.* Big Data: how BIG DATA technologies are changing our lives // Big Data and Advanced Analytics. 2021. No. 7-1. P. 189 (in Russ.).

[26] См.: *Roche J., Jamal A.* Op. cit. P. 340.

[27] См.: *Burtscher B., Fritz G.* (2015) Big data: who owns and who may use and exploit big data? URL: https://www.lexology.com/library/detail.aspx?g=77ab3ffb-8f25-469c-ad80-d14bbcee9551 (дата обращения: 04.01.2023).

4. *Andanda P.* Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research // IIC. 2019. Vol. 50. P. 1072, 1075.
5. *Bagnoli V.* Competition for the Effectiveness of Big Data Benefits // IIC – International Review of Intellectual Property and Competition Law. 2015. Vol. 46. P. 629–631.
6. *Burtscher B., Fritz G.* (2015) Big data: who owns and who may use and exploit big data? URL: https://www.lexology.com/library/detail.aspx?g=77ab3ffb-8f25-469c-ad80-d14bbcee9551 (accessed: 04.01.2023).
7. *Chirita A.D.* The Rise of Big Data and the Loss of Privacy // Personal Data in Competition, Consumer Protection and Intellectual Property Law / ed. by M. Bakhoum et al. Springer, 2018. P. 159, 175.
8. *De Bruin B.* The liberal value of privacy // Law and Philosophy. 2010. Vol. 29. P. 511, 532, 533.
9. DotEcon & Analysys Mason (2015), The Commercial Use of Consumer Data A research report for the CMA. URL: https://www.gov.uk/cma-cases/commercial-use-of-consumer-data (accessed: 04.01.2023).
10. *Fredriksson C., Mubarak F., Tuohimaa M., Zhan M.* Big Data in the Public Sector: A Systematic Literature Review // Scandinavian Journal of Public Administration. 2017. Vol. 21(3). P. 39–61.
11. *Froehlich A., Tăiatu C.M.* Space in Support of Human Rights. Cham, 2020. P. 36.
12. *Hosanagar K., Miller A.P.* Who Do We Blame for the Filter Bubble? On the Roles of Math, Data, and People in Algorithmic Social Systems // After the Digital Tornado: networks, algorithms, humanity / ed. by K. Werbach. Cambridge University Press, 2020. P. 111, 117, 118.
13. *Mittelstadt D.B., Floridi L.* The ethics of big data: current and foreseeable issues in biomedical contexts // Sci Eng Ethics. 2016. Vol. 22. P. 312.
14. *Rijmenam M.* Why the 3V's Are Not Sufficient To Describe Big Data // Datafloq. URL: https://datafloq.com/read/3vs-sufficient-describe-big-data/166 (accessed: 04.01.2023).
15. *Roche J., Jamal A.* A Systematic Literature Review of the Role of Ethics in Big Data // Cybersecurity, Privacy and Freedom Protection in the Connected World. Proceedings of the 13th International Conference on Global Security, Safety and Sustainability. London, January 2021 / ed. by H. Jahankhani, A. Jamal, S. Lawson. Cham, 2021. P. 337, 340.

Сведения об авторе

ТАЛАПИНА Эльвира Владимировна –
доктор юридических наук,
главный научный сотрудник Института
государства и права Российской академии наук;
119019 г. Москва, ул. Знаменка, д. 10

Authors' information

TALAPINA Elvira V. –
Doctor of Law,
Chief Researcher,
Institute of the State and Law
of the Russian Academy of Science;
10 Znamenka str., 119019 Moscow, Russia