

## ПРЕДЕЛЫ ОБРАБОТКИ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ ДЛЯ ЦЕЛЕЙ ПОЛУЧЕНИЯ ИНФОРМАЦИИ О ЧЕЛОВЕКЕ: ПРАВОВОЙ АСПЕКТ

**Гулемин Артем Николаевич**

Доцент кафедры информационного права Уральского государственного юридического университета им. В. Ф. Яковлева (Екатеринбург), кандидат юридических наук, ORCID: 0000-0003-0876-3849, e-mail: artem\_gulemin@mail.ru

*Исследуются общественные отношения по поводу анализа и обработки больших данных. Автор указывает на необходимость нормативного определения понятия больших пользовательских данных с учетом цели их сбора и обработки (создание новых сведений). Отмечается, что в результате хранения и обработки больших данных в единой информационной системе социальных сетей или цифровых экосистем для целей поведенческого таргетинга становится возможным создание новой информации о пользователях, которая может быть отнесена к специальным категориям персональных данных. Предлагается оценивать такую информацию с учетом признака достоверности и относить ее к персональным данным только после оформления согласия субъекта. Делается вывод о необходимости законодательного закрепления права субъекта получать информацию об алгоритмах обработки больших пользовательских данных в случае персонификации полученной в ходе нее информации. Подчеркивается важность соблюдения принципа соответствия содержания и объема данных заявляемым целям их автоматизированной обработки. Предлагается установить в законе обязанность оператора персональных данных обезличивать при автоматизированной обработке сами персональные данные и информацию, полученную в результате их обработки, а также хранить их в таком виде, чтобы они не могли быть персонифицированы в случае несанкционированного доступа к ним.*

*Ключевые слова: большие данные, персональные данные, обработка больших данных, информационные технологии, информационная безопасность, цифровая экономика*

*Для цитирования: Гулемин А. Н. Пределы обработки больших объемов данных для целей получения информации о человеке: правовой аспект // Электронное приложение к «Российскому юридическому журналу». 2022. № 6. С. 52–57. DOI: [https://doi.org/10.34076/22196838\\_2022\\_6\\_52](https://doi.org/10.34076/22196838_2022_6_52).*

## LIMITS OF BIG DATA PROCESSING FOR THE PURPOSES OF OBTAINING INFORMATION ABOUT A PERSON: A LEGAL ASPECT

**Gulemin Artem**

Associate professor, Ural State Law University named after V. F. Yakovlev (Yekaterinburg), candidate of legal sciences, ORCID: 0000-0003-0876-3849, e-mail: artem\_gulemin@mail.ru

*The article studies public relations regarding the analysis and processing of big data and the results of such processing. The author notes the need for a normative definition of the concept of big data, taking into account the purpose of their collection and processing (creation of new information). It is noted that storing and processing big data in a unified information system of social networks or digital ecosystems for behavioural targeting purposes make it possible to create new information about users, which can be classified as special categories of personal data. It is proposed to evaluate such information taking into account its reliability and attribute it to per-*

*sonal data only after the user's consent has been issued. The author concludes that it is necessary to legislate the subject's right to receive information about the algorithms for processing big user data in case of personification of the information received. He also underlines the need for observing the principle of compliance of the content and volume of data with the declared purposes of their automated processing. It is proposed to legislate the obligation of the operator of personal data to de-personalize the personal data themselves and the information obtained as a result of their processing. Also personal data should be stored in such a way that they cannot be personalized in case of unauthorized access.*

*Key words: big data, personal data, big data processing, information technology, information security, digital economy*

*For citation: 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 zhurnaluz»*, no. 6, pp. 52–57, DOI: [http://doi.org/10.34076/22196838\\_2022\\_6\\_52](http://doi.org/10.34076/22196838_2022_6_52).*

Большие данные – это новый феномен как в рамках технических, так и в рамках юридических наук. Исследования по вопросам обработки больших объемов информации осуществлялись еще с 90-х гг. XX в., однако само понятие больших данных начало использоваться только с 2005 г. в связи с появлением информационных технологий их обработки для целей управления бизнес-процессами и прогностической аналитики<sup>1</sup>. В настоящее время большие данные применяются практически во всех сферах экономической деятельности, во многих сквозных технологиях, включая индустриальный интернет, интернет вещей и искусственный интеллект. Цифровые технологии, основанные на анализе больших данных, все чаще используются компаниями с целью улучшения качества товаров и услуг и прогнозирования рыночных тенденций<sup>2</sup>.

В теории большие данные в зависимости от источника происхождения разделяются на промышленные большие данные – создаваемые без непосредственного участия человека для отображения состояния устройств индустриального интернета и интернета вещей – и большие пользовательские данные – создаваемые людьми в процессе использования различных приложений и сервисов в сети Интернет<sup>3</sup>. В рамках настоящего исследования интерес представляет именно вторая категория больших данных.

Попытки дать легальное определение понятию «большие данные» на национальном уровне до сих пор не увенчались успехом. Анализ единственного смежного понятия «обработка больших объемов данных», закрепленного в Стратегии развития информационного общества на 2017–2030 годы<sup>4</sup>, позволяет выделить существенные, по мнению законодателя, признаки больших данных:

это массив структурированной или неструктурированной информации (по инициативе заинтересованного лица собирается либо однородная информация, например данные о возрасте покупателей, либо любые сведения о лицах);

информация поступает из различных, в том числе не связанных между собой, источников (используются не одна, а несколько баз данных, например базы данных пользователей такси и покупателей магазина);

обработка таких данных невозможна вручную за разумное время.

Представляется, что законодатель упустил еще один важный признак больших данных – цель, для которой осуществляются их сбор и обработка. В случае законодательного определения понятия считаем необходимым включить в него формули-

<sup>1</sup> Войниканис Е. А. Регулирование больших данных и право интеллектуальной собственности: общие подходы, проблемы и перспективы развития // Закон. 2020. № 7. Доступ из СПС «КонсультантПлюс».

<sup>2</sup> Цариковский А. Ю., Иванов А. Ю., Войниканис Е. А. Антимонопольное регулирование в цифровую эпоху: как защищать конкуренцию в условиях глобализации и четвертой промышленной революции. М.: ВШЭ, 2018. С. 39.

<sup>3</sup> Савельев А. И. Направления регулирования Больших данных и защита неприкосновенности частной жизни в новых экономических реалиях // Закон. 2018. № 5. Доступ из СПС «КонсультантПлюс».

<sup>4</sup> Указ Президента РФ от 9 мая 2017 г. № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы».

ровку «целью сбора и обработки больших данных является создание новой информации». Создаваемые в процессе обработки сведения и выступают основой для принятия решений субъектами, осуществляющими обработку.

На первоначальном этапе формирования технологий обработки больших данных вопрос о том, из каких источников можно брать сведения о лицах, был решен в ставшем каноническим судебном споре социальной сети «ВКонтакте» и ООО «Дабл», касавшемся сбора пользовательских данных из общедоступных ресурсов для последующего определения кредитного рейтинга пользователей<sup>1</sup>. В этом споре Роскомнадзор четко обозначил, что данные в социальных сетях не являются общедоступными и их последующая обработка не может осуществляться без согласия субъекта<sup>2</sup>. Приведенная позиция многим исследователям кажется спорной<sup>3</sup>, и, видимо, поэтому она не нашла отражения в законодательстве.

Однако с появлением новых экономико-правовых субъектов – цифровых экосистем – данный вопрос приобретает «новое звучание». Внутри экосистемы пользователь может получить доступ ко всем сервисам через единый аккаунт, а обладателем разрозненных баз данных с информацией о совершаемых пользователем в экосистеме действиях становится одно лицо – бенефициар экосистемы.

Предполагается, что получаемые данные пользователей обрабатываются для целей поведенческого таргетинга (технология, позволяющая показывать пользователям рекламные объявления на основе их предпочтений). Процесс работы с большими данными в таком случае можно представить в виде последовательности действий: сбор данных о пользователе – их обезличивание (анонимизация) – обнаружение взаимосвязей между действиями пользователя – создание новой информации о нем – экстраполяция полученной информации на нового обезличенного пользователя. Однако возрастающие объемы обрабатываемых больших данных и использование все более совершенных, в том числе нейросетевых, технологий позволяют в настоящее время получить различного рода данные о конкретном (персонализированном, деанонимизированном) субъекте. На основе исходной информации о поездке человека в магазин на такси и покупке им определенных продуктов и лекарств можно установить, например, пять новых сведений: пол, возраст, социальный статус, религиозные и политические убеждения, состояние здоровья. Последовательное выявление этих сведений (хотя в реальности они могут обнаруживаться и непоследовательно) позволит с почти 60 %-ной вероятностью получить новые данные о субъекте, в том числе те, которые могут быть отнесены к специальным категориям персональных данных по смыслу ст. 10 ФЗ «О персональных данных»<sup>4</sup>. Исследования поведенческих стереотипов еще в 2013 г. позволяли с вероятностью от 0,85 до 0,95 определять расу, пол, политические предпочтения и вероисповедание пользователей по буквально трем «лайкам» в социальных сетях<sup>5</sup>.

Похожим образом развивается направление исследований геномной информации. Имея в своем распоряжении обезличенные «большие геномные данные» всего человечества, ученые могут установить причины многих заболеваний<sup>6</sup>. Однако риски утраты обезличенности могут привести к дискриминации и стигматизации человека, а дальнейшие разработки по сквозному геномному сканированию плода и новорожденного младенца обнажают угрозу установления евгенистической политики нации посредством реализации программ редактирования генома и отбора «экземпляров» с заданными характеристиками<sup>7</sup>.

<sup>1</sup> Орешин. Е. Вконтакте vs Дабл Дата как пример эффективного ведения судебного спора // Закон.ру. 2022. 24 сент. URL: [https://zakon.ru/blog/2022/9/24/vkontakte\\_vs\\_dabl\\_data\\_kak\\_primer\\_effektivnogo\\_vedeniya\\_sudebnogo\\_spora](https://zakon.ru/blog/2022/9/24/vkontakte_vs_dabl_data_kak_primer_effektivnogo_vedeniya_sudebnogo_spora) (дата обращения: 25.12.2022).

<sup>2</sup> Определение Верховного Суда РФ от 29 января 2018 г. № 305-КП17-21291 по делу № А40-5250/2017.

<sup>3</sup> См., например: Савельев А. И. Указ. соч.

<sup>4</sup> Федеральный закон от 27 июля 2006 г. № 152-ФЗ (ред. от 14 июля 2022 г.) «О персональных данных».

<sup>5</sup> Ледовая Я. А., Тихонов Р. В., Боголюбова О. Н. Социальные сети как новая среда для междисциплинарных исследований поведения человека // Вестник Санкт-Петербургского университета. Психология и педагогика. 2017. Т. 7. № 3. С. 201–203.

<sup>6</sup> Лисаченко А. В. Правовой режим «больших геномных данных»: за и против свободного обращения // Российский юридический журнал. 2022. № 2. С. 143.

<sup>7</sup> Митин А. Н., Кузнецов П. У. Организационно-правовые аспекты появления новых угроз от «наднационального контроля» в контексте информационной безопасности // Современное право. 2022. № 10.

Выше мы целенаправленно не использовали словосочетание «персональные данные», однако не потому, что большие данные – это деперсонифицированная информация, которая применяется заинтересованными компаниями в обезличенном виде<sup>1</sup>. Хотя законодатель и не устанавливает способы получения персональных данных и включает в их число любую информацию, относящуюся прямо или косвенно к определенному или определяемому физическому лицу, представляется, что полученную в результате обработки больших данных новую информацию о субъекте следует относить к персональным данным с учетом категории достоверности, которая, однако, также не определена в законодательстве. Как бы ни развивались технологии обработки информации, обеспечение стопроцентной точности выявляемых сведений кажется невозможным. На наш взгляд, единственным способом установления достоверности информации о человеке, полученной после обработки больших данных, является оформленное согласие субъекта с выводами, которые были сделаны компьютером.

При использовании новой информации, полученной после обработки больших данных, возникает два практических вопроса.

Первый – принятие юридически значимых решений после автоматизированной обработки персональных данных. Считается, что этот вопрос урегулирован законодательством. Статья 16 ФЗ «О персональных данных» запрещает принятие исключительно на основании автоматизированной обработки персональных данных решений, порождающих юридические последствия в отношении субъекта персональных данных или иным образом затрагивающих его права и законные интересы, однако устанавливает ряд исключений: если такое решение принято при наличии согласия в письменной форме субъекта персональных данных или в случаях, предусмотренных федеральными законами. По смыслу статьи решение принимается именно машиной на основании скрытых в ней алгоритмов. Однако используемая в ней формулировка может ввести в заблуждение, поскольку наводит на мысль, что решение принимает человек.

Закон устанавливает обязанность оператора разъяснить субъекту персональных данных порядок принятия решения на основании исключительно автоматизированной обработки его персональных данных и возможные юридические последствия такого решения, предоставить возможность заявить возражение против такого решения, а также разъяснить порядок защиты субъектом персональных данных своих прав и законных интересов. Однако исследователи отмечают, что эти положения «касаются не того, как именно принято решение, к примеру, что положено в его основу, а, скорее, процедуры его принятия. Из указанного следует, что субъект персональных данных, несмотря на то что наделен правом подать возражение на решение, фактически им воспользоваться не сможет, поскольку не будет знать, как именно было принято решение. В то же время основополагающей задачей правового регулирования должно являться создание условий, при которых человек имеет право на получение исчерпывающей информации относительно логики принятого в его отношении решения»<sup>2</sup>.

В условиях существования технической возможности получения, например, новых сведений о состоянии здоровья человека на основе его покупок в аптеке и последующего автоматизированного принятия решения об отказе в выдаче кредита появляется необходимость законодательно закрепить возможность субъекта получить информацию о причинно-следственной связи этих двух событий (разрешить техническую проблему «черного ящика» при принятии решения). Требуется также обеспечить соблюдение принципа соответствия содержания и объема обрабатываемых данных заявляемым целям их автоматизированной обработки. Для этого нужно признать недопустимым получение на основе обработки больших данных новых сведений, которые после персонификации могут быть отнесены к специальным категориям персональных данных или биометрическим персональным данным субъекта. Со-

<sup>1</sup> Кириллова Е. А. Правовой статус и принципы использования технологии больших данных (Big Data) // Российская юстиция. 2021. № 2. С. 68–69.

<sup>2</sup> Кутейников Д. Л., Ижаев О. А., Зенин С. С., Лебедев В. А. Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика» // Lex russica (Русский закон). 2020. Т. 73. № 6. С. 146.

гласие пользователя на обработку его персональных данных алгоритмами обработки больших данных должно выражаться осознанно, субъект должен действительно ознакомиться с условиями пользовательского соглашения, изложенными в доступном для восприятия виде<sup>1</sup>.

Второй вопрос – несанкционированное использование больших данных или результатов их обработки вследствие хищения или иных противоправных действий злоумышленника и, как результат, вмешательство в частную жизнь субъекта. Похищение персональных данных, избыточное профилирование сужают личное пространство человека и нарушают его приватность. Раскрытие личной информации «стало неотъемлемым атрибутом корыстных преступлений и преступлений против личности»<sup>2</sup>. В последнее время в СМИ все чаще пишут о фактах незаконного получения, обнародования и распространения персональных данных. Зачастую такие утечки происходят в силу того, что операторы персональных данных недостаточно обеспечивают информационную безопасность в рамках своей организации. Вызывают сомнение факты привлечения организаций, допустивших утечку персональных данных, к ответственности по ч. 1 ст. 13.11 КоАП РФ. Считаем необходимым законодательно установить обязанность любого оператора обезличивать при автоматизированной обработке персональных данные, а также сведения, полученные в результате обработки больших данных, которые в дальнейшем могут быть отнесены к персональным данным, и хранить их в таком виде, чтобы они не могли быть персонифицированы в случае несанкционированного доступа к ним.

Изложенное позволяет сделать следующие выводы:

1. Категория больших данных должна быть раскрыта в законодательстве не только через их технологические, но и формально-юридические признаки, включая цель сбора и обработки больших данных – создание новой значимой информации.

2. Полученная в ходе обработки больших пользовательских данных новая информация должна оцениваться с учетом ее достоверности и признаваться одним из видов персональных данных только после оформления согласия субъекта с выводами, сделанными в результате такой обработки.

3. Следует законодательно закрепить возможность субъекта ознакомиться с алгоритмами, на основе которых была получена новая информация о нем, а также в соответствии с принципами ФЗ «О персональных данных» ограничить обработку больших данных в случаях, если итогом такой обработки будет получение сведений, которые в дальнейшем могут быть отнесены к специальным категориям персональных данных или биометрическим персональным данным.

4. Необходимо законодательно закрепить обязанность оператора персональных данных обезличивать при автоматизированной обработке сами персональные данные и информацию, полученную в результате их обработки, а также хранить их в таком виде, чтобы они не могли быть персонифицированы в случае несанкционированного доступа к ним.

### Список литературы

Войниканис Е. А. Регулирование больших данных и право интеллектуальной собственности: общие подходы, проблемы и перспективы развития // Закон. 2020. № 7. С. 135–156.

Кириллова Е. А. Правовой статус и принципы использования технологии больших данных (Big Data) // Российская юстиция. 2021. № 2. С. 68–69.

Кутейников Д. Л., Ижаев О. А., Зенин С. С., Лебедев В. А. Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика» // Lex russica (Русский закон). 2020. Т. 73. № 6. С. 139–148.

Ледовая Я. А., Тихонов Р. В., Боголюбова О. Н. Социальные сети как новая среда для междисциплинарных исследований поведения человека // Вестник Санкт-Петербургского университета. Психология и педагогика. 2017. Т. 7. № 3. С. 193–210.

Лисаченко А. В. Правовой режим «больших геномных данных»: за и против свободного обращения // Российский юридический журнал. 2022. № 2. С. 140–151.

<sup>1</sup> Сергеев А. П., Терещенко Т. А. Большие данные: в поисках места в системе гражданского права // Закон. 2018. № 11. Доступ из СПС «КонсультантПлюс».

<sup>2</sup> Паршуков М. И. Преследование в информационном пространстве: проблемы понятийного аппарата, опыт правового регулирования государств – участников Содружества Независимых Государств // Вестник ЮУрГУ. Сер.: Право. 2021. Т. 21. № 1. С. 110.

Митин А. Н., Кузнецов П. У. Организационно-правовые аспекты появления новых угроз от «наднационального контроля» в контексте информационной безопасности // Современное право. 2022. № 10. С. 35–47.

Орешин. Е. Вконтакте vs Дабл Дата как пример эффективного ведения судебного спора // Закон.ру. 2022. 24 сент. URL: [https://zakon.ru/blog/2022/9/24/vkontakte\\_vs\\_dabl\\_data\\_kak\\_primer\\_effektivnogo\\_vedeniya\\_sudebnogo\\_spora](https://zakon.ru/blog/2022/9/24/vkontakte_vs_dabl_data_kak_primer_effektivnogo_vedeniya_sudebnogo_spora) (дата обращения: 25.12.2022).

Паршуков М. И. Преследование в информационном пространстве: проблемы понятийного аппарата, опыт правового регулирования государств – участников Содружества Независимых Государств // Вестник ЮУрГУ. Сер.: Право. 2021. Т. 21. № 1. С. 109–115.

Савельев А. И. Направления регулирования Больших данных и защита неприкосновенности частной жизни в новых экономических реалиях // Закон. 2018. № 5. С. 122–144.

Сергеев А. П., Терещенко Т. А. Большие данные: в поисках места в системе гражданского права // Закон. 2018. № 11. С. 106–123.

Цариковский А. Ю., Иванов А. Ю., Войниканис Е. А. Антимонопольное регулирование в цифровую эпоху: как защищать конкуренцию в условиях глобализации и четвертой промышленной революции. М.: ВШЭ, 2018. 311 с.

## References

Kirilova E. A. (2021) Pravovoi status i printsipy ispol'zovaniya tekhnologii bol'shikh dannykh (Big data) [Legal status and principles of using big data technology]. In *Rossiiskaya yustitsiya*, no. 2, pp. 68–69.

Kuteinikov D. L., Izhaev O. A., Zenin S. S., Lebedev V. A. (2020) Algoritmicheskaya prozrachnost' i podotchetnost': pravovye podkhody k razresheniyu problemy «chernogo yashchika» [Algorithmic transparency and accountability: legal approaches to solving the black box problem]. In *Lex russica*, vol. 73, no. 6, pp. 139–148.

Ledovaya Ya. A., Tikhonov R. V., Bogolyubova O. N. (2017) Sotsial'nye seti kak novaya sreda dlya mezhdistsiplinarnykh issledovaniy povedeniya cheloveka [Social networks as a new environment for interdisciplinary research of human behavior]. In *Vestnik Sankt-Peterburgskogo universiteta. Psikhologiya i pedagogika*, vol. 7, no. 3, pp. 193–210.

Lisachenko A.V. (2022) Pravovoi rezhim «bol'shikh genomnykh dannykh»: za i protiv svobodnogo obrashcheniya [The legal regime of «big genomic data»: pros and cons of free circulation]. In *Rossiiskii yuridicheskii zhurnal*, no. 2, pp. 140–151.

Mitin A. N., Kuznetsov P. U. (2022) Organizatsionno-pravovye aspekty poyavleniya novykh ugroz ot «nadnatsional'nogo kontrolya» v kontekste informatsionnoi bezopasnosti [Organizational and legal aspects of the emergence of new threats from «supranational control» in the context of information security]. In *Sovremennoe pravo*, no. 10, pp. 35–47.

Oreshin E. (2022) Vkontakte vs Dabl Data kak primer effektivnogo vedeniya sudebnogo spora [Vkontakte vs Double Data as an example of effective litigation]. In *Zakon.ru*, 24 Sept., available at: [https://zakon.ru/blog/2022/9/24/vkontakte\\_vs\\_dabl\\_data\\_kak\\_primer\\_effektivnogo\\_vedeniya\\_sudebnogo\\_spora](https://zakon.ru/blog/2022/9/24/vkontakte_vs_dabl_data_kak_primer_effektivnogo_vedeniya_sudebnogo_spora) (accessed: 25.12.2022).

Parshukov M. I. (2021) Presledovanie v informatsionnom prostranstve: problemy ponyatiinogo apparata, opyt pravovogo regulirovaniya gosudarstv – uchastnikov Sodruzhestva Nezavisimykh Gosudarstv [Persecution in the information space: problems of the conceptual apparatus, experience of legal regulation of the member states of the Commonwealth of Independent States]. In *Vestnik YuUrGU. Ser.: Pravo*, vol. 21, no. 1, pp. 109–115.

Savel'ev A. I. (2018) Napravleniya regulirovaniya Bol'shikh dannykh i zashchita neprikosnovennosti chastnoi zhizni v novykh ekonomicheskikh realiyakh [Directions of Big Data regulation and protection of privacy in the new economic realities]. In *Zakon*, no. 5, pp. 122–144.

Sergeev A. P., Tereshchenko T. A. (2018) Bol'shie dannye: v poiskakh mesta v sisteme grazhdanskogo prava [Big data: in search of a place in the system of civil law]. In *Zakon*, no. 11, pp. 106–123.

Tsarikovskii A. Yu., Ivanov A. Yu., Voinikanis E. A. (2018) *Antimonopol'noe regulirovanie v tsifrovuyu epokhu: kak zashchishchat' konkurentsiyu v usloviyakh globalizatsii i chetvertoi promyshlennoi revolyutsii* [Antitrust regulation in the digital age: how to protect competition in the face of globalization and the fourth industrial revolution]. Moscow, VShE, 311 p.

Voynikanis E. A. (2020) Regulirovanie bol'shikh dannykh i pravo intellektual'noi sobstvennosti: obshchie podkhody, problemy i perspektivy razvitiya [Big Data regulation and intellectual property law: general approaches, problems and development prospects]. In *Zakon*, no. 7, pp. 135–156.