

Шайдуллина Венера Камилевна

кандидат юридических наук, директор Центра исследований и экспертиз Департамента правового регулирования экономической деятельности Финансового университета при Правительстве Российской Федерации

БОЛЬШИЕ ДАННЫЕ И ЗАЩИТА ПЕРСОНАЛЬНЫХ ДАННЫХ: ОСНОВНЫЕ ПРОБЛЕМЫ ТЕОРИИ И ПРАКТИКИ ПРАВОВОГО РЕГУЛИРОВАНИЯ**Аннотация:**

Настоящая статья является результатом исследования, посвященного анализу вопросов применения российского законодательства о защите персональной информации в условиях развития больших данных (big data). В работе автор раскрывает указанное понятие, описывает генезис и преимущества этой технологии, изучает аналитику по проблемам масштабов распространения больших данных в разных отраслях экономики. Определено, что формируемые большими данными возможности находятся в прямом противоречии с международными принципами защиты персональных сведений и ставят под сомнение эффективность и адекватность действующего отечественного законодательства о личной информации. Рассмотрен вопрос, относятся ли депersonализованные данные к персональным. Установлена важность переосмысления таких категорий, как «персональные данные», «обезличенные данные» и «оператор персональных данных», в рамках российского законодательства.

Ключевые слова:

большие данные, персональные данные, big data, информационные брокеры, профайлинг, обезличивание.

Shaydullina Venera Kamilevna

PhD in Law, Director of the Center for Research and Expertise, Department of Legal Regulation of Economic Activity, Financial University under the Government of the Russian Federation

BIG DATA AND PERSONAL DATA PROTECTION: THE MAIN THEORETICAL AND PRACTICAL ISSUES OF LEGAL REGULATION**Summary:**

The paper analyzes the application of Russian legislation on personal data protection in the context of big data development. The study reveals the concept of big data, describes the genesis and benefits of this technology, and examines big data distribution analytics in various sectors of the economy. The research identifies that the opportunities generated by big data are in direct conflict with international principles of personal data protection and call into question the effectiveness and adequacy of the current Russian legislation on personal data. The author considers whether depersonalized data are personal. The researcher concludes that it is necessary to rethink such categories as personal data, anonymized data, and personal data processor within the framework of Russian laws.

Keywords:

big data, personal data, information brokers, profiling, depersonalization.

Развитие современных информационных технологий привело к преобразованию нашего понимания личного пространства и частной жизни. Процессы, ранее происходившие в физическом (реальном) мире, перетекли в онлайн-среду: приобретение услуг и товаров, общение со знакомыми и друзьями, взаимодействие с органами государства, работодателями и др. Вследствие этого объемы личных сведений, которые человек раскрывает о себе и выкладывает в глобальную сеть, как и персональных данных граждан, собираемых и систематизируемых разными учреждениями и ведомствами, увеличились до беспрецедентных размеров. Такая ситуация сложилась в связи с действием значительного числа факторов, а именно: 1) проникновения интернета во все сферы жизни; 2) развития электронной торговли; 3) возникновения и совершенствования поисковых сервисов, в основе которых находится рекламная бизнес-модель, предполагающая сбор огромного массива данных о поведении интернет-пользователей; 4) появления соцсетей, агрегирующих информацию об индивидах и отношениях между ними; 5) повсеместного распространения планшетов и смартфонов, дающих людям возможность постоянно находиться онлайн, обмениваться мгновенными сообщениями и отслеживать маршрут передвижения пользователей.

Все перечисленные факторы поспособствовали тому, что основные процессы человеческой жизнедеятельности перетекли в интернет. Сегодня любые действия оставляют цифровой след, что приводит к возникновению значительных массивов виртуальных данных. Статистика объема информации, создаваемой в мире в течение последних лет, поражает воображение. В 2017 г. количество хранящейся информации составляло 16,2 зеттабайта, из которых на нецифровые данные приходилось меньше 3 % [1, р. 12]. По прогнозам специалистов компании IDC,

которая специализируется на аналитике в области ИТ, общий объем данных будет увеличиваться вдвое каждые 2 года и к 2025 г. составит около 163,0 зеттабайта [2]. Большинство сведений, произведенных в период 2012–2025 гг., будут сгенерированы не людьми, а разными устройствами в ходе их взаимодействия между собой и сетями данных (например, смартфонами, сенсорами, системами спутниковой навигации (GPS, ГЛОНАСС) и др.) [3].

Взаимодействие устройств через интернет предполагает их цифровую идентификацию, привязанную к базам данных. Оно заложено в основу концепции «интернета вещей», рассматриваемого как следующий этап эволюции развития Всемирной сети, где машины являются не только производителями информации, но и ее потребителями. На современном этапе общество характеризуется резким увеличением объемов данных, циркулирующих в нем, а также устойчивой тенденцией к возрастанию роли информации в разных сферах деятельности (включая предпринимательскую и отдельные области государственного управления).

Рассмотренные тенденции – стремительное повышение объема данных, циркулирующих во всем мире, и их безусловная коммерческая ценность – предъявляют ряд новых требований к механизмам обработки информации, а также к извлечению из нее добавленной стоимости. Ответом на этот вызов стали технологии, которые получили в бизнес-среде и технической сфере обобщенное наименование «большие данные» (big data). Востребованность последних подтверждается существенным расширением соответствующего сектора рынка.

В Стратегии развития отрасли информационных технологий в РФ на период 2014–2020 гг. и на перспективу до 2025 г. большие данные указываются в числе технологий, являющихся прорывными для мировой индустрии, которые в ближайшие 10–15 лет с высокой вероятностью могут обеспечить глобальную технологическую конкурентоспособность РФ [4]. Кроме того, Правительство РФ утвердило программу «Цифровая экономика Российской Федерации», где указывается, что развитие сквозных данных (в том числе больших данных) является одним из направлений цифровизации [5].

Однако до настоящего времени так и не сложилось общепринятого определения термина «большие данные». Наиболее распространено раскрытие этого феномена через указание проблем, возникающих при обработке сведений [6]. Большие данные предполагают возможность высокопроизводительного анализа информации, способного обеспечить работу в онлайн-режиме. Рассматриваемая методология предполагает исследование ситуации, сложившейся на текущем этапе, для того чтобы повлиять на нее [7]. Таким образом, big data – это совокупность методов и инструментов обработки огромных объемов структурированной и неструктурированной информации из разных источников для повышения качества принятия управленческих решений, увеличения конкурентоспособности и создания новых продуктов [8].

Первые специальные положения, регулирующие автоматизированную обработку персональных данных, появились в Европе и затем распространились по всему миру. По данным исследования Финансового университета при Правительстве Российской Федерации, в 2017 г. законы о персональных данных имеются в 104 странах мира.

Основополагающий акт в этой сфере – Конвенция от 28.01.1981 г. о защите физических лиц при автоматизированной обработке персональных данных, принятая Советом Европы. На основе конвенционных положений на национальном уровне странами Евросоюза были приняты отдельные законы, посвященные регулированию использования персональных сведений. В дальнейшем национальное законодательство было гармонизировано несколькими директивами ЕС. В настоящее время на территории Европейского союза действует Общий регламент по защите данных (Regulation (EU) 2016/679 (EU GDPR)).

В 2005 г. Россией была ратифицирована Конвенция СЕ от 1981 г. о защите физических лиц при автоматизированной обработке персональных данных. В результате этого был принят закон № 152-ФЗ от 27.07.2006 г. «О персональных данных» [9]. Этот акт отражает принципы защиты и обработки персональных сведений, принятые в европейских странах.

Следует отметить, что создаваемые большими данными возможности находятся в прямом противоречии с конвенционными принципами и ставят под сомнение эффективность и адекватность действующего законодательства о персональных сведениях. Некоторые исследователи сделали выводы, что большие данные и право на частную жизнь несовместимы [10].

Особый акцент в эпоху больших данных делается на повторном использовании информации, обладающей огромной ценностью. Сведения о совершенных клиентами покупках позволяют настраивать адресную рекламу, строить прогнозы по поводу статуса покупателя и проводить рыночные исследования [11].

Современные технологии (в том числе облачных вычислений) поспособствовали устранению многих ограничений, присущих сбору данных. Сбор и хранение огромных массивов информации стали доступны и недороги. В эпоху больших данных компании заинтересованы в получении как можно более существенного объема сведений для хранения и дальнейшего использования. Учитывая современное развитие технологий, проследить выполнение законодательных требований на практике будет довольно трудно.

Идеальной моделью регламентации вопросов защиты персональных данных, на наш взгляд, является формирование такого законодательства, в соответствии с которым субъект персональных данных вправе самостоятельно принимать решения, затрагивающие его информационную сферу, путем согласия на обработку его персональных сведений. Такое согласие – главное легитимирующее основание подобной обработки. Чтобы оно было сознательным, конкретным и информированным, необходимо обеспечить предоставление субъекту детальных сведений о том, каким образом, в каких целях и какими способами будут использоваться его данные (ч. 4 ст. 9, ч. 7 ст. 14 Федерального закона «О персональных данных» от 27.07.2006 г. № 152-ФЗ).

В эпоху big data концепция информированного согласия на обработку персональных сведений не обладает должной эффективностью, и тому есть несколько причин: невозможность предоставления исчерпывающего объема информации о целях и способах обработки данных; неспособность субъекта персональных данных к адекватному восприятию такой информации; невозможность индивидуально взаимодействовать с огромным числом компаний, занимающихся сбором и обработкой личных сведений. Все это приводит к неутешительному выводу: информированное согласие обеспечивает принятие субъектом ряда дискретных решений на ранних этапах обработки (обычно на стадии сбора) персональных данных, но в силу специфики применения технологий big data предугадать последствия подобных решений нельзя. В результате информированное согласие является не более чем фикцией и не играет роли основного легитимирующего основания для обработки персональной информации.

К числу мер, направленных на минимизацию риска причинения вреда гражданам в случае утечки, относится обезличивание персональных данных. Это действия, в результате которых становится невозможным без использования дополнительных сведений определить принадлежность личной информации конкретному человеку (ст. 3 № 152-ФЗ). Нормы в указанной сфере приняты Приказом Роскомнадзора от 05.09.2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных». Основными способами анонимизации выступают введение идентификаторов, перемешивание, изменение семантики или состава, декомпозиция. Оператор самостоятельно выбирает способ, исходя из задач и целей обработки данных. При этом нужно учитывать, что обезличивание должно обеспечивать не только защиту личной информации от неправомерного применения, но также возможность ее полноценной обработки. Другими словами, после обезличивания данные должны обладать некоторыми свойствами, конкретный набор которых будет зависеть от технологии обезличивания. К таким свойствам относятся полнота и анонимный характер информации.

Встает вопрос: относятся ли деперсонализованные данные к персональным либо это отдельный вид данных, не подпадающий под режим персональных? Широко распространено мнение, что деперсонализация выводит информацию из-под режима персональной, что является вполне удобной альтернативой необходимости соблюдать обременительные нормы, связанные с ее обработкой [12]. По мнению Европейской рабочей группы, если анонимизированные данные обратимы, т. е. могут быть возвращены в исходное состояние, то они входят в категорию сведений, которые способствуют опосредованному определению субъекта, а следовательно – относятся к числу персональных [13, р. 18].

В Законе о персональных данных прямого ответа на этот вопрос не содержится. С учетом того что дефиниции персональных данных в праве РФ и стран Европы близки, есть все основания полагать, что обезличенные сведения имеют статус персональных с определенной спецификой. В частности, в указанном документе предусмотрены два специальных правила для обезличенной информации:

- 1) возможность обработки данных в исследовательских и статистических целях без получения согласия со стороны пользователя (п. 9 ч. 1 ст. 9);
- 2) обезличивание данных как альтернатива удалению персональной информации после достижения целей обработки (ч. 7 ст. 5) [14].

Поскольку многими методами деперсонализации, указанными в приказе Роскомнадзора, предполагается обратимость, имеются основания считать, что деперсонализация не выводит личную информацию за рамки действия № 152-ФЗ. Исключением является только изменение семантики или состава. Тем не менее анонимизация персональных сведений имеет большие перспективы, так как это средство, по мнению некоторых исследователей, служит максимально эффективным способом решения вопроса о защите персональной информации в условиях ее повсеместной передачи в сети Интернет. Однако в эпоху big data, когда существует возможность идентификации личности путем установления корреляций между несколькими фрагментами данных, действенность этого метода сомнительна. Причем неважно, какой из механизмов обезличивания используется [15].

Таким образом, техническая возможность деанонимизации обусловлена двумя ключевыми факторами: увеличением производительности и доступности вычислительных мощностей, а также огромным массивом имеющихся в Мировой сети личных данных. В эпоху big data сведения могут или быть анонимными, или представлять ценность для обработки, но одновременно обладать обеими характеристиками они не могут. Чем выше степень обезличивания данных, тем меньше их ценность для анализа. Если на современном этапе информация действительно представляет новую «нефть», то наивно ожидать, что организации будут ее уничтожать, вместо того чтобы извлекать выгоду [16].

Однако сказанное не означает, что обезличивание персональных сведений бесполезно и от него стоит отказаться. Речь идет о том, что его не следует рассматривать как средство, безусловно достаточное для эффективной защиты личной информации в эпоху больших данных [17].

После возникновения технологий big data понятие «информационное общество» обрело первоначальный смысл, а информация получила статус ценнейшего актива, выступающего в качестве движущей силы информационного общества. Большие данные могут эффективно использоваться в разных сферах: банковской и страховой деятельности, медицине, электронной торговле и правоохранительной работе. Тем не менее у них имеются и недостатки, главный из которых – вторжение в частную жизнь людей. Также следует отметить, что технологии big data обнажают очевидный факт: действующее законодательство о персональных данных все меньше и меньше отвечает требованиям современности, в связи с чем нуждается в скорейшем реформировании.

Должны быть переосмыслены категории «персональные данные», «оператор персональных данных». В настоящее время, когда сбор информации о пользователях обладает массовым характером, даже самый безобидный фрагмент подобных сведений, будучи соединенным с другими, способен дать гораздо больше данных о человеке, чем его анкета. Нужно ли придавать таким единицам статус персональной информации или нет смысла в их выделении в особую категорию с отдельным регулированием – эти вопросы требуют незамедлительного разрешения.

Значительные массивы сведений о пользователях, циркулирующих в цифровой форме, привели к возникновению за рубежом новых игроков на информационном рынке. В частности, появились информационные брокеры, которые на основе стекающихся к ним из разных интернет-сервисов данных составляют детальные профайлы людей и предоставляют к ним доступ заинтересованным субъектам. Существующее регулирование, выстроенное на понятии оператора персональных данных, не принимает во внимание значительные различия между следующими лицами:

- 1) осуществляющими первичный сбор информации (интернет-сервисы);
- 2) на профессиональной основе агрегирующими такие данные в профайлы посредством технологий big data;
- 3) приобретающими доступ к подобным профайлам для собственных потребностей (к примеру, финансовые структуры могут покупать эти сведения для оценки платежеспособности клиентов).

По сути, все перечисленные лица – операторы персональных данных. Однако совершаемые ими действия носят различный характер и могут повлечь разные последствия для лица, предоставившего личные сведения, с позиции возможного вреда вследствие неправомерной обработки.

Указанные проблемы невозможно решить до тех пор, пока не будет найден ответ на вопрос о том, что представляют собой персональные данные – товар или неотчуждаемое неимущественное благо. В пользу каждого варианта имеются аргументы, однако приведенные рассуждения позволяют склоняться к первому варианту. Ведь важно то, что большинство эффективных бизнес-схем в интернете основаны на применении персональных данных как своеобразной «валюты», которой пользователи расплачиваются за доступ к соответствующим сервисам. Именно сведения о пользователе – источник многомиллиардных доходов поисковых ресурсов, социальных сетей и разных ИТ-компаний. Однако этот факт игнорируют не только многие пользователи, но и отечественные законодатели, которые не признают предоставление согласия на обработку персональных данных в качестве встречного предоставления в целях квалификации соглашения как возмездного (ст. 424 ГК РФ). Следует подчеркнуть, что от ответа на этот вопрос напрямую зависит и преодоление всех остальных противоречий, обозначенных ранее.

Ссылки:

1. Reinsel D., Gantz J., Rydning J. The Digitization of the World. From Edge to Core [Электронный ресурс] : IDC Report // Seagate Technology. 2018. 28 p. URL: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> (дата обращения: 29.01.2019).
2. Gantz J., Reinsel D. The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East [Электронный ресурс] // IDC. 2012. Dec. URL: <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf> (дата обращения: 29.01.2019).

3. White T. Hadoop: The Definitive Guide [Электронный ресурс] // O'Reilly Media, Inc. 2012. 3rd ed. 647 p. URL: <https://www.isical.ac.in/~acmsc/WBDA2015/slides/hg/Oreilly.Hadoop.The.Definitive.Guide.3rd.Edition.Jan.2012.pdf> (дата обращения: 29.01.2019).
4. Об утверждении Стратегии развития отрасли информационных технологий в РФ на 2014–2020 гг. и на перспективу до 2025 г. [Электронный ресурс] : распоряжение Правительства РФ от 1 нояб. 2013 г. № 2036-р : в ред. от 18 окт. 2018 г. Доступ из справ.-правовой системы «КонсультантПлюс».
5. Об утверждении программы «Цифровая экономика РФ» [Электронный ресурс] : распоряжение Правительства РФ от 28 июля 2017 г. № 1632-р. Доступ из справ.-правовой системы «КонсультантПлюс».
6. Laney D. 3D Data Management: Controlling Data Volume, Velocity, and Variety [Электронный ресурс] : Application Delivery Strategies // META Group. 2001. Febr. 6. URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (дата обращения: 29.01.2019).
7. Ramanathan S. Data to Big Data – A Paradigm Shift and a Professional Challenge [Электронный ресурс] // CSI Communications. 2014. July. P. 36–37. URL: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3305A447A27E06092C5E76E7CFAD2975?doi=10.1.1.588.5588&rep=rep1&type=pdf> (дата обращения: 29.01.2019).
8. Савельев А.И. Проблемы применения законодательства о персональных данных в эпоху «больших данных» (big data) // Право. Журнал Высшей школы экономики. 2015. № 1. С. 43–66.
9. О персональных данных [Электронный ресурс] : федер. закон от 27 июля 2006 г. № 152-ФЗ. Доступ из справ.-правовой системы «КонсультантПлюс».
10. Шилина М.Г. Big data и цифровая датификация как техносциальный феномен. К вопросу формирования научно-теоретической рамки исследования // Социально-политические науки. 2018. № 4. С. 60–65.
11. Соколова А.С. Влияние технологий анализа «больших данных» (big data) на законодательство о персональных данных // Юриспруденция 2.0: новый взгляд на право : материалы межвузовской научно-практической конференции с международным участием. М., 2017. С. 282–285.
12. Поддымникова А.Е., Воробьева К.Р., Мастилин А.Е. Большие данные. Расширение области практического применения // Экономика и социум. 2018. № 5 (48). С. 993–995.
13. Opinion 4/2007 on the Concept of Personal Data WP 136 [Электронный ресурс] // European Commission. 2007. 26 p. URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf (дата обращения: 29.01.2019).
14. Соколова А.С. Указ. соч.
15. Карчагин Е.В. Справедливость в информационной области: этический аспект big data // Вестник Волгоградского государственного университета. Сер. 7: Философия. Социология и социальные технологии. 2016. № 1 (31). С. 50–58.
16. Володенков С.В. Технологии big data в современных политических процессах: цифровые вызовы и угрозы // Вестник Томского государственного университета. Философия. Социология. Политология. 2018. № 44. С. 205–212.
17. Артемьева И.Н. Большие данные – большие деньги. Facebook как источник и продавец big data // Смирновские чтения – 2018. Цифровая экономика и финансовые кибертехнологии: проблемы и перспективы : материалы XVII Международной научно-практической конференции. СПб., 2018. С. 27–33.

References:

- Artemyeva, IN 2018, 'Big Data is Big Money. Facebook as a Source and Seller of Big Data', *Smirnovskiye chteniya – 2018. Tsvirovaya ekonomika i finansovyye kibertekhnologii: problemy i perspektivy: materialy XVII Mezhdunarodnoy nauchno-prakticheskoy konferentsii*, St. Petersburg, pp. 27-33, (in Russian).
- Gantz, J & Reinsel, D 2012, 'The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East', *IDC*, Dec., viewed 29 January 2019, <<http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf>>.
- Karchagin, EV 2016, 'Information Justice: the Ethical Aspect of Big Data', *Vestnik Volgogradskogo gosudarstvennogo universiteta. Ser. 7: Filosofiya. Sotsiologiya i sotsial'nyye tekhnologii*, no. 1 (31), pp. 50-58, (in Russian).
- Laney, D 2001, '3D Data Management: Controlling Data Volume, Velocity, and Variety: Application Delivery Strategies', *META Group*, Febr. 6, viewed 29 January 2019, <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>>.
- 'Opinion 4/2007 on the Concept of Personal Data WP 136' 2007, *European Commission*, 26 p., viewed 29 January 2019, <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>.
- Poddymnikova, AE, Vorobyeva, KR & Mastilin, AE 2018, 'Big Data. Expanding the Field of Practical Application', *Ekonomika i sotsium*, no. 5 (48), pp. 993-995, (in Russian).
- Ramanathan, S 2014, 'Data to Big Data – A Paradigm Shift and a Professional Challenge', *CSI Communications*, July, pp. 36-37, viewed 29 January 2019, <<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3305A447A27E06092C5E76E7CFAD2975?doi=10.1.1.588.5588&rep=rep1&type=pdf>>.
- Reinsel, D, Gantz, J & Rydning, J 2018, 'The Digitization of the World. From Edge to Core: IDC Report', *Seagate Technology*, 28 p., viewed 29 January 2019, <<https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-data-age-whitepaper.pdf>>.
- Saveliev, AI 2015, 'Problems of Legislation Application on Personal Data in the Era of Big Data', *Pravo. Zhurnal Vyshey shkoly ekonomiki*, no. 1, pp. 43-66, (in Russian).
- Shilina, MG 2018, 'Big Data and Digital Certification as a Technosocial Phenomenon. To the Question of the Formation of the Scientific and Theoretical Research Framework', *Sotsial'no-politicheskiye nauki*, no. 4, pp. 60-65, (in Russian).
- Sokolova, AS 2017, 'The Impact of Big Data Analysis Technologies on Big Data Legislation', *Yurisprudentsiya 2.0: novyy vzglyad na pravo: materialy mezhvuzovskoy nauchno-prakticheskoy konferentsii s mezhdunarodnym uchastiyem*, Moscow, pp. 282-285, (in Russian).
- Volodenkov, SV 2018, 'Big Data Technologies in Modern Political Processes: Digital Challenges and Threats', *Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya*, no. 44, pp. 205-212, (in Russian).
- White, T 2012, 'Hadoop: The Definitive Guide', *O'Reilly Media*, 3rd ed., 647 p., viewed 29 January 2019, <<https://www.isical.ac.in/~acmsc/WBDA2015/slides/hg/Oreilly.Hadoop.The.Definitive.Guide.3rd.Edition.Jan.2012.pdf>>.