ПРОБЛЕМЫ ПРИМЕНЕНИЯ ЗАКОНОДАТЕЛЬСТВА О ПЕРСОНАЛЬНЫХ ДАННЫХ В ЭПОХУ «БОЛЬШИХ ДАННЫХ» (BIG DATA)*

А.И. Савельев

старший научный сотрудник научно-учебной лаборатории по информационному праву Национального исследовательского университета «Высшая школа экономики», юрисконсульт IBM Россия/СНГ, член Консультативного Совета при Роскомнадзоре, кандидат юридических наук. Адрес: 101000, Российская Федерация, Москва, ул. Мясницкая, 20. E-mail: garantus@rambler.ru

Аннотация

Настоящая статья представляет собой одно из первых исследований в России, посвященных анализу влияния технологий анализа «Больших данных» (Big Data) на законодательство о персональных данных, которое выступает одним из основных гарантов защиты права граждан на неприкосновенность частной жизни в цифровой среде. В статье раскрываются понятие «Больших данных», описывается генезис данной технологии и ее преимущества, а также приводятся примеры реализации данной технологии в различных сферах деятельности. Основное внимание уделяется анализу совместимости «Больших данных» с рядом базовых положений законодательства о персональных данных. По результатам анализа делается вывод, что такие принципы, как ограничение обработки персональных данных заранее определенными целями, ограничения объема собираемых и обрабатываемых данных минимально необходимым объемом, осуществление обработки данных на основе информированного согласия являются несовместимыми с природой технологий «Больших данных», которая лежит в основе тех преимуществ, которые она несет в себе. Так, принципы ограничения обработки персональных данных заранее определенными целями и ограничения объема обрабатываемых данных минимально необходимым объемом несовместимы с идеей повторного использования данных, которой пронизана философия «Больших данных». Информированное согласие невозможно в условиях, когда невозможно указание цели обработки персональных данных, а оно невозможно как раз по причине непредсказуемости таких целей в эпоху “Больших данных”: ограничение обработки персональных данных заранее определенными целями означает лишение данной технологии преимуществ, которые она способна предоставить. При этом решение проблемы посредством популяризации обезличивания персональных данных также не может оправдать возлагаемых на него надежд по причине существования широких возможностей по деобезличиванию таких данных, предоставляемых дешевыми вычислительными мощностями и большими массивами общедоступных данных в сети Интернет.

Ключевые слова

персональные данные, Большие данные, Big Data, обезличивание, профайлинг, информационные брокеры.

* Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ в 2015 году.

Библиографическое описание: 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.

JEL: K39.

1. Введение

1.1. Предпосылки появления «Больших данных»

Современные информационные технологии в значительной степени преобразили наше понимание частной жизни и личного пространства. Те процессы, которые ранее происходили в реальном (физическом) мире, перетекли в онлайновую среду: приобретение товаров и услуг, общение с друзьями и знакомыми, взаимодействие с работодателем, государственными органами и т.д. Как следствие, объемы личной информации, которые лицо раскрывает о себе и выкладывает в сеть Интернет и, соответственно, объемы личных данных граждан, подвергаемых сбору и систематизации различными органами и организациями, возросли до беспрецедентных размеров. В значительной степени такое положение вещей сложилось в силу одновременного действия множества факторов: 1) проникновения Интернета в повседневную жизнь; 2) развития электронной коммерции; 3) появления и развития поисковых сервисов, имеющих в своей основе рекламную бизнес-модель, предполагающую сбор огромных массивов информации о поведении индивидов в сети Интернет; 4) появления социальных сетей, которые агрегируют данные не только об индивидах, но и о отношениях между ними; 5) повсеместного распространения смартфонов и планшетов, позволяющих быть постоянно онлайн, отслеживать маршрут передвижения своих пользователей, а также обмениваться мгновенными сообщениями. Как следствие, ключевые процессы жизнедеятельности человека перетекли в Интернет и любое действие индивида оставляет цифровой след, что в совокупности повлекло появление огромных размеров массива цифровой информации.

Статистика объема данных, создаваемых на протяжении последних лет, поражает воображение. В 2013 г. количество хранящейся в мире информации составило 1,2 зеттабайта (около 1,2 млн. петабайт или 1,2 трлн. гигабайт), из которых на нецифровую информацию приходится менее 2%[1]. По прогнозам компании IDC, специализирующейся на аналитике в сфере информационных технологий, общее количество информации будет удваиваться каждые 2 года и составит к 2020 г. порядка 40 зеттабайт[2]. При этом бóльшая часть данных, которая будет произведена в период с 2012 по 2020 годы, будет сгенерирована не людьми, а различного рода устройствами в ходе их взаимодействия друг с другом и сетями данных (например, сенсорами, смартфонами, устройствами радиочастотной идентификации (RFID), спутниковыми системами навигации типа ГЛОНАСС или GPSи т.д.)[3]. Взаимодействие различного рода устройств между собой посредством сети Интернет, предполагающее их цифровую идентификацию, привязан-

[1] *Майер-Шенбергер В., Кукьер К.* Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. М. 2014. С. 17.

[2] *Gantz J.* and *Reinsel D.* The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. December 2012. URL: http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf (дата доступа: 15.09.2014)

[3] *White T.* Hadoop: The Definitive Guide. O'Reilly Media, Inc. 3rd ed., 2012. P. 2.

ную к базам данных, лежит в основе концепции «Интернета вещей» (Internet of Things)[4], который рассматривается в качестве следующего этапа эволюции развития сети Интернет, где машины будут являться не только производителями информации, но и ее потребителями.

Однако современное информационное общество характеризуется не только стремительным ростом объемов информации, циркулирующей в нем. Помимо этого существует устойчивая тенденция к возрастанию роли информации в различных сферах деятельности, начиная от предпринимательской и заканчивая отдельными сферами государственного управления. Информация все больше приобретает статус нового фактора производства, который нередко и не без оснований называют «новой нефтью»[5]. В частности, обладание большими массивами информации в совокупности с необходимыми инструментами для ее сбора и обработки позволяет создавать товары и услуги с высокой долей добавленной стоимости, а также принимать эффективные управленческие решения[6]. Исследования, проведенные Массачусетским технологическим университетом (MIT) показали, что организации, использующие в своей деятельности автоматизированные механизмы принятия решений, основанные на анализе данных, увеличивали свою производительность в среднем на 5-6%[7]. Эксперты отмечают, что «сбор, интеграция и анализ данных больше не считаются расходами на ведение бизнеса; данные — это ключ к достижению эффективности и прибыльности бизнеса. В результате быстро развивается индустрия, поддерживающая анализ данных»[8].

Обозначенные тенденции — стремительный рост информации, циркулирующей по всему миру и ее очевидная коммерческая ценность — предъявляют новые требования к технологиям обработки данных и извлечения из них добавленной стоимости. Ответом на этот вызов стали технологии, получившие в технической и бизнес-среде обобщенное название «Большие данные» (*Big Data*). Востребованность данной технологии иллюстрируется стремительным ростом соответствующего рынка. Так, согласно прогнозу IDC, вышедшему в марте 2012 года, рынок технологий и сервисов для обработки «Больших данных» вырастет с $3,2 млрд. в 2010 году до $16,9 млрд. в 2015 году. Это соответствует среднегодовому темпу роста (CAGR) на уровне 40%, что примерно в 7 раз больше, чем среднегодовой темп роста всего рынка информационных техноло-

---

[4] Считается, что термин «Интернет вещей» был первоначально использован еще в 1999 г. Кевином Эштоном (Kevin Ashton), ученым MIT, который стоял у истоков технологии RFID. Он отмечал, что на первых порах развития сети Интернет, составляющая его информация была создана людьми. На следующих этапах развития сети бóльшая часть информации, наполняющей ее, будет создана устройствами, что позволит оперативно решать вопросы, связанные с их обслуживанием, обладая информацией о том, когда вышло из строя, требует замены или усовершенствования. В конечном итоге он приходит к выводу, что Интернет вещей может изменить мир в той же степени, как это сделал в свое время сам Интернет. *Ashton K.* That 'Internet of Things' Thing // RFID Journal. July 22, 2009.

[5] Как отмечается, «Данные — это новая «нефть»: они представляют собой ценность, однако в необработанном состоянии они не могут быть использованы». *Arthur C.* Tech giants may be huge, but nothing matches big data // The Guardian, August 23, 2013. URL: http://www.theguardian.com/technology/2013/aug/23/tech-giants-data(дата обращения: 15.09.2014)

[6] *Uden L., Aho A.-M.* Knowledge Management in Organizations. 9th International Conference, KMO 2014. Springer International. P. 99.

[7] *Brynjolfsson E.* et al. Strength in Numbers: How Does Data-Driven Decision -Making Affect Firm Performance? (April 22, 2011). Available at SSRN: http://ssrn.com/abstract=1819486(дата обращения: 15.09.2014).

[8] *Rakesh A.* et al. The Claremont Report on Database Research. 2008. URL: http://db.cs.berkeley.edu/claremont/claremontreport08.pdf (дата обращения: 15.09.2014)

гий в целом[9]. В Стратегии развития отрасли информационных технологий в Российской Федерации на 2014–2020 годы и на перспективу до 2025 года технологии обработки «Больших данных» обозначены в числе «прорывных для мировой индустрии, в которых в перспективе 10–15 лет с высокой вероятностью может быть обеспечена глобальная технологическая конкурентоспособность России»[10].

1.2. Понятие технологий «Больших данных»

Термин «Большие данные» (Big Data) не имеет общепринятого определения даже в индустрии информационных технологий. Наиболее распространенным является раскрытие феномена «Больших данных» через указание проблем, с которыми приходится сталкиваться на современном этапе развития технологий при обработке информации. Исходя из этого «Большие данные» определяются посредством указания следующих основных характеристик: 1) большого объема *(Volume)*, 2) разнообразия данных (*Variety)* и 3) высокой скорости их изменения *(Velocity)*. Данный подход получил название «трех «V»[11].

Согласно указанному подходу, помимо собственно обработки больших объемов данных (*Volume*) проблема, решаемая посредством Big Data, состоит также и в том, что бóльшая часть потенциально ценной информации представлена в неструктурированном виде, то есть не упорядочена и содержится в различных форматах, в отличие от данных, которые наполняют традиционные базы данных (*Variety*). Огромные массивы разнообразной информации, например, информация с форумов и социальных сетей, видеозаписи, текстовые документы, лог-файлы или, например, данные о трафике и соединениях абонентов, содержатся в различных источниках, нередко за пределами организации. В результате компании могут иметь доступ к огромному объему данных из внутренних и внешних источников и не иметь необходимых инструментов, чтобы осуществить их совместную обработку, выявив определенные взаимосвязи и сделать на их основе значимые выводы. Технологии «Больших данных» позволяют решить эту проблему, связав воедино разнородные данные. Если же данные достаточно единообразны и структурированы, то есть их можно легко разбить по строкам и столбцам традиционной базы данных, то несмотря на большой объем, для их анализа вполне подходят уже имеющиеся методы: принципиально новых технологий обработки данных не требуется, достаточно увеличения производительности уже имеющихся.

Третий признак «Больших данных» (*velocity*) состоит в том, что обрабатываемая с использованием указанной технологии информация обновляется быстро (например, «потоковые данные» — streaming data), при этом необходимо принимать решения на основании их оперативного анализа. Традиционные подходы к анализу информации не могут угнаться за огромными объемами постоянно обновляемых данных[12]. Мето-

[9] IDC. Worldwide Big Data Technology and Services 2012–2015 Forecast. March 2012. URL: http://www.idc.com/research/viewtoc.jsp?containerId=233485 (дата обращения: 15.09.2014).

[10] Распоряжение Правительства РФ от 01.11.2013 N 2036-р «Об утверждении Стратегии развития отрасли информационных технологий в Российской Федерации на 2014 — 2020 годы и на перспективу до 2025 года»

[11] *Laney D.* 3-D Data Management: Controlling Data Volume, Velocity and Variety // Application Delivery Strategies. META Group. February 6, 2001. URL: http://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/ (дата обращения: 15.09.2014).

[12] В качестве иллюстрации роли оперативности в производстве анализа соответствующих данных можно привести следующую аналогию: нельзя безопасно перейти дорогу, если в руке у пешехода фото-

дология «Больших данных» предполагает возможность высокопроизводительного анализа данных, обеспечивающего обработку информации в режиме реального времени (например, за счет выполнения вычислений непосредственно в оперативной памяти компьютера). Как отмечается, если при традиционном подходе к анализу данных (подготовка отчетов на основе накопленных данных), предпринималась попытка проанализировать прошлое, чтобы совершить действие в будущем, феномен «Больших данных» предполагает анализ настоящего, чтобы повлиять на текущую ситуацию[13].

Таким образом, «Большие данные» можно определить как **совокупность инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов из различных источников, подверженных постоянным обновлениям, в целях повышения качества принятия управленческих решений, создания новых продуктов и повышения конкурентоспособности.** Консалтинговая компания «Форрестер» дает краткую формулировку: «Большие данные» объединяют техники и технологии, которые извлекают смысл из данных на экстремальном пределе практичности»[14].

2. Примеры возможных направлений применения технологий «Больших данных» в различных сферах деятельности

Технологии Big Data носят универсальный характер и могут быть использованы в самых различных сферах деятельности. Применительно к коммерческим их видам основной интерес технологии “Больших данных” будут представлять для организаций, где уже накоплены большие массивы данных по клиентам и операционной деятельности. Особенно интересны для иллюстрации возможных направлений применения рассматриваемой технологии в контексте эволюции отношений с клиентами организации финансового сектора (банки и страховые организации), организации, функционирующие в сфере электронной коммерции и розничной продажи («ритейл»). Немалую роль технологии «Больших данных» могут сыграть и в медицине, а также различных аспектах правоохранительной деятельности государства.

2.1. Банковский сектор

В сфере банковской деятельности посредством технологий Big Data может осуществляться анализ платежеспособности потенциального заемщика или лица, предоставляющего обеспечение по займу. В США активно развивается индустрия, получившая название информационных брокеров (Data Brokers), которые агрегируют сведения о миллионах граждан США и ряда других стран из самых разнообразных источников (общедоступная информация в социальных сетях; данные о транзакциях и займах, полу-

графия того, как выглядело дорожное движение на ней 5 минут назад. Для принятия решения в данном случае необходима наиболее актуальная информация.

[13] *Ramanathan S.* Data to Big Data — A Paradigm Shift and a Professional Challenge // CSI Communications. July 2014. P. 36; *Davenport T., Barth P., Bean R.* How ‘Big Data’ Is Different // MIT Sloan Management Review 54, no. 1 (Fall 2012). URL:http://sloanreview.mit.edu/article/how-big-data-is-different/(дата обращения: 15.09.2014).

[14] *Uden L., Aho A.-M.* Op. cit. P. 100.

ченные от партнеров, о правонарушениях, налоговых выплатах и т.д.), систематизируют их по категориям граждан, например: «находящиеся на грани выживания», «на пенсии без накоплений», «родители-одиночки», «неплатежеспособные семьи, проживающие в городе» и др. Доступ к таким данным предоставляется на условиях подписки заинтересованным лицам, в том числе кредитным учреждениям[15]. Очевидно, обладание подобного рода сведениями нередко позволяет делать более точные выводы относительно платежеспособности потенциального заемщика, по сравнению с данными, которые могут храниться в бюро кредитных историй[16] и даже делать прогнозы о том, как платежеспособность будет меняться в будущем. В свою очередь это помогает не только более эффективно управлять кредитным портфелем, но и значительно упрощать процедуру выдачи кредита за счет сокращения количества необходимых для предъявления документов (например, справок о доходах, сведений о составе семьи и иждивенцах и т.д.).

Информация о клиентах и их действиях, совершенных за определенный промежуток времени, позволяет выработать индивидуализированный подход к клиенту, в частности, выдвигая более адресные и направленные предложения дополнительных услуг. Так, знание истории трат клиентов позволяет предсказывать, на что они потратят деньги в будущем. Обладая таким прогнозом, можно предлагать каждому клиенту нужный именно ему продукт в нужный момент — как свой банковский (например, выгодный кредит на планируемую покупку), так и предложение партнёров. Это очевидным образом увеличивает кросс-продажи и некомиссионные доходы банка, а также повышает лояльность клиента: банк из обезличенного инструмента превращается в персонального финансового помощника[17]. В качестве иллюстрации можно привести опыт сингапурского подразделения City Bank, которое по данным о транзакциях клиентов, их локации и времени суток, в которое они были проведены, делало выводы о вкусах клиента и направляло ему индивидуальное предложение. Например, если известно, что клиент любит итальянскую кухню и в обеденное время расплатился в такси банковской карточкой City Bank рядом с улицей, где есть итальянский ресторан, с которым у банка заключено партнёрское соглашение, клиент получает смс-уведомление от мобильного клиента банка со специальным предложением в этом заведении[18].

С другой стороны, подобного рода индивидуализация может принимать весьма неожиданные формы. В частности, были случаи, когда компания American Express использовала данные клиента о совершенных им покупках для изменения размера кредитного лимита по его кредитной карте. В качестве обоснования при этом использовались сведения, что «другие клиенты, которые использовали карты для расчетов в указанных местах, продемонстрировали низкий уровень платежной дисциплины»[19]. Таким обра-

---

[15] A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes. Committee on Commerce, Science and Transportation. Staff Report for Chairman Rockefeller. US Senate. December 18, 2013. P. ii. URL: http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=0d2b3642-6221-4888-a631-08f2f255b577 (дата обращения: 15.09.2014)

[16] В России такие данные, как правило, ограничены лишь сведениями о выданных займах (кредитах) и ходе их погашения, либо отказах в выдаче займа (кредита). См.: ст. 4 Федерального закона от 30.12.2004 N 218-ФЗ «О кредитных историях» (в ред. от 28.06.2014), которая содержит перечень информации, включающейся в кредитную историю.

[17] *Блинов Н.* Банки и большие данные — знать о клиенте больше, чем он знает о себе сам // Future Banking. 21.05.2014. URL: http://futurebanking.ru/post/2511 (дата обращения: 15.09.2014).

[18] Там же.

[19] *Сиото С.* et al. 'GMA' Gets Answers: Some Credit Card Companies Financially Profiling Customers // ABC News. 28.01.2009. URL: http://goo.gl/YPjtTR (Датаобращения: 14.02.2015).

зом, совершение покупок в магазинах дисконтов, неожиданное прекращение покупок товаров, которые приобретались ранее на регулярной основе — все это может служить аргументом в пользу вывода об ухудшении финансового положения соответствующего клиента и повлечь определенную реакцию со стороны банка.

2.2. Страховая деятельность

Страховая деятельность, по заявлению главы компании Google Эрика Шмидта, является наиболее очевидной областью применения технологий Больших данных[20]. Это следует из самой сути осуществляемой ими деятельности: необходимости анализа вероятности наступления страхового случая для оценки принимаемых на себя страховых рисков и определения адекватного размера страховых взносов. Как отмечал М.И. Брагинский, «ключевое значение при подсчете размера страховой премии имеют так называемые актуарные расчеты. Под ними подразумевается совокупность экономико-математических методов расчетов необходимого и достаточного объема ресурсов страхового фонда страховщика. В основе актуарных расчетов лежит использование действия закона больших чисел»[21].

Так, например, крупная британская страховая компания Aviva вместо прохождения потенциальными страховщиками жизни и здоровья процедур сдачи анализов стала использовать кредитные отчеты и данные потребительского маркетинга, позволившие ей выявить лиц, наиболее подверженных риску развития высокого артериального давления, диабета или депрессии. Этот метод основывался на данных об образе жизни страховщика, включающих сотни переменных (хобби, посещаемые веб-сайты и время, затрачиваемое на просмотр телевизора и т.д.). Или другой пример. Благодаря распространению в Великобритании беспроводных модулей, помещенных в транспортные средства, водители могут приобрести автостраховку, стоимость которой определяется не только исходя из стандартных параметров вроде возраста, пола или даты последней аварии, но и времени и маршрута фактических поездок страхователя[22]. В России подобный подход был недавно внедрен страховой компанией Intouch, которая предложила своим клиентам бесплатно устанавливать в автомашины специальный модуль от компании МТС[23].

2.3. Электронная коммерция

Продажи крупных торговых компаний вроде Amazon.com, Inc. (крупнейшего в мире интернет-магазина) и Wal-Mart Stores, Inc. (крупнейшей в мире розничной сети) построены на работе с «Большими данными». Некоторые компании создают собственные инструменты или даже лаборатории, которые фокусируются на изучении поведения пользователей.

---

[20] *Womack B., Trish R.* Google's Schmidt: Insurance About to 'Explode' With Uses for Big Data // Insurance Journal. November 25,2013. URL:http://www.insurancejournal.com/news/national/2013/11/25/312031.htm(Дата обращения: 14.02.2015).

[21] *Брагинский М.И., Витрянский В.В.* Договорное право. Т. 3: Договоры о выполнении работ и оказании услуг. М.: Статут, 2011. С. 567.

[22] *Майер-Шенбергер В., Кукьер К.* Указ. соч. С. 96.

[23] *Королев И.* МТС занялась «умным страхованием» автомобилистов // CNews. 19.09.2014. URL:http://bigdata.cnews.ru/top/2014/09/19/mts_zanyalas_umnym_strahovaniem_avtomobilistov_586504 (дата обращения: 19.09.2014)

Автоматизированная система рекомендаций от Amazon определяет товары, способные заинтересовать покупателя, на основе оценок, которые он ставил на веб-сайте, и покупок, которые ранее совершил. Таким образом, чем больше книг или других товаров клиент заказал, тем лучше алгоритм понимает его потребности и тем более точную выборку товаров предлагает. Аналогичные подходы применяет и российский аналог Amazon — интернет-магазин Ozon.ru.

В целом простор для использования технологий Big Data в сфере электронной коммерции весьма велик. В частности, это анализ поведения покупателей на веб-сайте магазина: их виртуального маршрута и продолжительности визита, случаев незавершенных покупок. На основе выявленных характеристик неконкурентоспособных товаров (цена, качество, доставка, цвет) в совокупности со сведениями из профиля клиента в социальных сетях (количество друзей, количество подписчиков, «вес» на графе связей в социальных сетях, частота сообщений), компании могут в реальном времени выделить наиболее обсуждаемые товары, повысить степень удовлетворенности покупателей и получить более широкий охват аудитории в сети Интернет.

Хрестоматийным считается пример использования технологий «Больших данных» американской сетью магазинов Target, которая внедрила в процесс взаимодействия с клиентами результаты автоматизированной аналитики данных, накопленных компанией за несколько лет, в частности, сведений о транзакциях по банковским и именным скидочным картам. Соответствующие алгоритмы проанализировали, как и в каких условиях менялись предпочтения покупателей и на основе сгенерированных прогнозов покупателям делали всевозможные специальные предложения. Весной 2012 года разразился скандал, когда отец двенадцатилетней школьницы пожаловался, что его дочери присылают буклеты с предложениями для беременных. Первоначально сеть Target была готова признать ошибку и извиниться перед покупателем, однако вскоре выяснилось, что девочка действительно была беременна, хотя ни она, ни ее отец на момент жалобы не знали об этом. Но алгоритм уловил изменения в поведении покупательницы, характерные для беременных женщин[24].

2.4. Медицинская сфера

Технологии «Больших данных» позволяют обеспечить индивидуальный подход к лечению пациента. В частности, устройства, использующие технологии «Больших данных», могут выступить в качестве персонального помощника врача. На основе анализа данных о передовых научных исследованиях в соответствующей сфере, имеющемся опыте лечения соответствующих заболеваний, клинических исследований отдельных лекарственных препаратов, может быть разработан индивидуальный план лечения для пациента, учитывающий особенности его организма. Необходимость «внедрения технологий масштабирования баз знаний и внедрения систем поддержки принятия врачебных решений в повседневную деятельность» была отмечена в государственной программе РФ «Развитие здравоохранения» 2014 г. Ряд компаний уже активно работает в данном направлении. Например, компания IBM в настоящее время сотрудничает с рядом организаций в области здравоохранения для разработки систем, которые смогут предоставлять результаты анализа генома человека и сократить время, необходимое для подбора правильного лечения пациента. Они будут собирать информацию о геноме

[24] *Kashmir H*. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did // Forbes. 16.02.2012. URL: http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/ (дата обращения: 14.02.2015).

и реакции пациентов на прописанное лечение и предлагать варианты, подобранные с учетом особенностей ДНК конкретного человека[25].

2.5. Большие данные в сфере профилактики правонарушений

Технологии «Больших данных» могут применяться и в инновационных подходах к профилактике правонарушений. В частности, широкую известность получила система Blue CRUSH (от англ.: Crime Reduction Utilizing Statistical History — «Снижение преступности на основе статистических данных»), разработанная компанией IBM, которая поставляет полицейским подготовленные на основе имеющейся статистики совершения преступлений сведения о зонах потенциальной угрозы совершения преступления с указанием места (в пределах нескольких кварталов) и времени (в пределах нескольких часов конкретного дня недели)[26]. Подобного рода профилактическое прогнозирование привело к снижению уровня преступности в г. Мемфисе на 31%, из которых 15% приходится на тяжкие преступления[27]. Технические решения, имеющие в своей основе аналитику «Больших данных», используются в ряде иных городов США (Нью-Йорк, Сиэтл, Лос-Анджелес и др.) и масштаб их использования возрастает с каждым годом[28].

Большие данные используются в качестве одного из ключевых компонентов программ массовой слежки за гражданами, осуществляемыми Агентством национальной безопасности США (АНБ) в отношении как граждан США, так и иностранных лиц. По данным газеты «Гардиан» от 2013 года, ежедневно системы сбора информации АНБ перехватывали и записывали около 1,7 млрд. телефонных разговоров и электронных сообщений и около 5 млрд. записей о местонахождении и передвижениях владельцев мобильных телефонов по всему миру[29]. При этом основным источником данных являлись американские компании Microsoft, Google, Yahoo, Facebook, America Online и Apple, предоставлявших АНБ прямой доступ к своим серверам[30]. В основе программы PRISM, используемой АНБ США для слежки, и считающейся одной из наиболее эффективных применяются те же компоненты, что и в «общегражданских», традиционных решениях Big Data (например, программное обеспечение с открытым исходным кодом Hadoop, о котором будет подробнее сказано далее[31]). Тот факт, что АНБ осуществляет деятельность в масштабах «Больших данных» с использованием передовых технологий, подтверждают и сами представители АНБ[32].

---

[25] *Wullianallur R.*, *Viju R.* Big Data Analytics in Healthcare: Promise and Potential // Health Information Science and Systems. No. 2. Vol. 3. 2014, URL:http://www.hissjournal.com/content/2/1/3

[26] http://www.ibm.com/smarterplanet/us/en/leadership/memphispd/(Дата обращения: 14.02.2015).

[27] *Thompson T.* Crime Software May Help Police Predict Violent Offences // The Guardian, July 25, 2010. URL: http://www.theguardian.com/uk/2010/jul/25/police-software-crime-prediction (Дата обращения: 14.02.2015).

[28] *Joh E.* Policing by Numbers: Big Data and the Fourth Amendment // Washington Law Review No. 89:35. 2014. P. 35-68.

[29] *Glenn G.* Are all Telephone Calls Recorded and Accessible to the US Government? A Former FBI Counterterrorism Agent Claims on CNN that this is the Case // The Guardian. May 4, 2013.

[30] *Gellman B.*, *Poitras L.* U.S., British Intelligence Mining Data from Nine U.S. Internet Companies in Broad Secret Program // The Washington Post. June 7, 2013 URL: http://goo.gl/KoTn4z (Дата обращения: 14.02.2015).

[31] *Harris D.* Under the Covers of the NSA's Big Data Effort // Gigaom. June 7, 2013. URL: https://gigaom.com/2013/06/07/under-the-covers-of-the-nsas-big-data-effort/(Дата обращения: 14.02.2015).

[32] *Rajesh D.* The NSA and Accountability in an Era of Big Data // Journal of National Security Law & Policy. No. 7, 2014. P. 310.

Как видно из вышеуказанных примеров, технологии «Больших данных» сулят государству и отдельно взятым компаниям большие выгоды, от которых могут выиграть и обычные граждане. Однако при этом степень вмешательства в частную жизнь отдельно взятого гражданина может быть высокой, причем чем больше персональных данных о лице агрегируется и подвергается обработке, тем выше степень возможного влияния на жизнь такого лица результатов автоматизированной обработки и, соответственно, величина риска, связанная с нарушением его прав.

Основным законодательным барьером, стоящим на пути возможных злоупотреблений в сфере обработки больших массивов информации о гражданах, является законодательство о персональных данных. До появления информационных технологий сбор, обработка и хранение персональных данных были крайне дорогостоящим занятием как для компаний, так и для государства, что служило своего рода «естественным барьером» личного пространства физического лиц[33]. Появление возможности автоматизированной обработки таких данных в значительной степени снизили его значение, что обусловило появление альтернативного «правового барьера», который бы позволил защитить личное пространство физического лица.

Однако, как будет показано далее, несмотря на относительную эффективность положений законодательства о персональных данных применительно к устоявшимся методам обработки массивов данных, изолированных рамками отдельных организаций, технологии «Больших данных» несовместимы с рядом базовых принципов, лежащих в основе законодательства о персональных данных, что обуславливает необходимость его реформирования.

3. «Большие данные» и законодательство о персональных данных

Специальные положения, посвященные проблематике автоматизированной обработки персональных данных, сначала появились в Европе и впоследствии распространились по всему миру. По состоянию на начало 2012 г. законы о персональных данных были приняты в 89 странах мира[34].

Основополагающим актом в данной сфере стала Конвенция о защите физических лиц при автоматизированной обработке персональных данных, принятая Советом Европы 28 января 1981 г., впоследствии дополненная протоколом по вопросам полномочий наблюдательных органов и трансграничной передачи данных. На основе положений данной Конвенции на национальном уровне страны Европы приняли отдельные законы, посвященные регулированию персональных данных. Впоследствии национальное законодательство было гармонизировано рядом директив ЕС, в числе которых следует

[33] *Войниканис Е.* Право интеллектуальной собственности в цифровую эпоху. Парадигма баланса и гибкости. М., 2013. С. 199. Схожие соображения были высказаны Верховным судом США в решении по делу United States v. Jones (132 S. Ct., 963, 2012): «В до-компьютерную эпоху лучшая защита тайны частной жизни предоставлялась не законом, а существующими реалиями. Традиционные средства сбора информации о гражданах в течение продолжительного срока были сопряжены с рядом сложностей и затратами, в силу чего редко имели место быть на практике»

[34] *Greenleaf G.* Global Data Privacy Laws: 89 Countries, and Accelerating // Privacy Laws & Business International Report, N 115, February 2012; Queen Mary School of Law Legal Studies Research Paper No. 98/2012. URL: http://ssrn.com/abstract=2000034 (дата обращения: 14.02.2015).

упомянуть Директиву 95/46/ЕС от 24 октября 1995 г. о защите прав частных лиц применительно к обработке персональных данных и о свободном движении таких данных[35] и Директиву 2002/58/ЕС от 31 июля 2002 г.[36], касающуюся использования персональных данных и защиты неприкосновенности частной жизни в сфере телекоммуникаций. Указанные директивы были имплементированы в национальное законодательство государств — членов ЕС.

В 2005 г. Россия ратифицировала Конвенцию Совета Европы о защите физических лиц при автоматизированной обработке персональных данных 1981 г.[37], в результате чего был принят ФЗ от 27 июля 2006 г. N 152-ФЗ «О персональных данных», вступивший в силу 26 января 2007 г. (далее — Закон о персональных данных).

Указанный закон в значительной степени отражает принципы защиты и обработки персональных данных, которые приняты в Европе и считаются основополагающими[38]. Среди них ст. 5 Закона о персональных данных предусматривает:

1) законное основание обработки персональных данных, в качестве которого выступает информированное, конкретное и сознательное согласие субъекта персональных данных, за исключением случаев, прямо указанных в законе;

2) ограничение обработки персональных данных только заранее установленными, законными целями; не допускается обработка персональных данных, несовместимая с целями их сбора, а равно избыточность персональных данных по отношению к целям их обработки;

3) недопустимость объединения баз данных, содержащих персональные данные, обработка которых осуществляется в целях, несовместимых между собой;

4) при обработке персональных данных должна быть обеспечена их точность и актуальность по отношению к целям обработки персональных данных. Оператор должен принимать необходимые меры либо обеспечивать их принятие по удалению или уточнению неполных или неточных данных.

5) хранение персональных данных должно осуществляться в форме, позволяющей определить субъекта персональных данных не дольше, чем этого требуют цели обработки персональных данных.

Как будет показано далее, возможности, создаваемые технологиями «Больших данных», находятся в прямом противоречии с указанными принципами и в целом ставят под сомнение адекватность и эффективность законодательства о персональных данных в его нынешнем виде применительно к новейшим технологическим реалиям. Некоторые исследователи уже делают категорические выводы о том, что право на частную жизнь и «Большие данные» несовместимы между собой[39].

---

[35] Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data

[36] Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector [as amended by Directive 2009/136/EC]

[37] ФЗ от 19.12.2005 г. N 160-ФЗ "О ратификации Конвенции Совета Европы о защите физических лиц при автоматизированной обработке персональных данных". Процесс ратификации был завершен 15 мая 2013 г. Конвенция вступила в силу в отношении России с 1 сентября 2013 г.

[38] *Carey P.* Data Protection: A Practical Guide to UK and EU Law. Oxford University Press, 2004. P. 51-65; *Bygrave L.* Data Privacy Law: An International Perspective. Oxford University Press. 2014. P. 153.

[39] Lane J., Stodden V., Bender S., Nissenbaum H. Privacy, Big Data, and the Public Good: Frameworks for Engagement. Cambridge: Cambridge University Press. 2014. P. 70.

3.1. «Большие данные» несовместимы с принципом ограничения обработки персональных данных заранее определенными целями

Принцип определенности целей сбора и обработки данных является одним из основополагающих, будучи предопределенным правом индивида на сообщение ограниченного перечня сведений о себе в строгом соответствии с конкретной необходимостью[40]. В эпоху «Больших данных» главный акцент делается на повторном использовании данных, поскольку *все* без исключения данные приобретают потенциальную ценность. Это в равной степени относится как к техническим данным (показателям датчиков температуры на заводе), так и к данным, потенциально связанным с конкретными пользователями (данные GPSнавигаторов, данные о посещении определенных сайтов, оставленных комментариях, сделанных поисковых запросах или покупках), то есть той информации, которая может быть потенциально квалифицирована как персональные данные, дефиниция которых является предельно широкой («любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу» — ст. 3 Закона о персональных данных).

Такого рода повторное использование данных имеет немалую ценность. Данные о сделанных клиентами покупках позволяют не только делать адресную рекламу, но и проводить исследования рынка, а также строить прогнозы относительно статуса клиента (пример с компанией Target, определившей беременность покупательницы по характеру ее покупок). Чем больше данных у компании, тем больше у нее простора для применения технологий «Больших данных» в целях выявления различного рода закономерностей, которые могут иметь значение для принятия бизнес-решений.

В литературе о «Больших данных» подчеркивается, что «истинная ценность данных — как айсберг в океане. На первый взгляд видна лишь незначительная их часть, в то время как все остальное сокрыто под водой. Инновационные компании, которые понимают это, могут извлечь скрытую ценность и получить огромные преимущества»[41]. Такая скрытая ценность может быть зачастую получена путем объединения одного набора данных с другим, на первый взгляд, совершенно с ним не связанным, поскольку при анализе «Больших данных» совокупность важнее отдельных частей, а при перекомпоновке совокупностей нескольких наборов данных получается еще более удачная совокупность. Существуют специальные Интернет-сервисы, получившие название «мэшапов» (от англ. mash-up), которые по-новому объединяют несколько источников данных.

Современные технологии также устранили большинство ограничений, которые были присущи сбору данных: запись и хранение огромных массивов данных стала доступной и недорогой, немалая доля заслуги в чем принадлежит технологиям облачных вычислений. Поскольку стоимость хранения упала, оправдать сбор и хранение огромных массивов информации стало гораздо проще, что стимулирует менеджмент организаций к принятию прагматичных решений об игнорировании принципа ограничения обработки персональных данных заранее определенной целью, равно как и ряд иных положений законодательства о персональных данных.

---

[40] *Бачило И.Л., Сергиенко Л.А., Кристальный Б.В., Арешев А.Г.* Персональные данные в структуре информационных ресурсов. Основы правового регулирования. Минск. 2006. С. 23.

[41] *Майер-Шенбергер В., Кукьер К.* Указ. соч. С. 111.

Таким образом, в эпоху "Больших данных" организации заинтересованы в том, чтобы собирать как можно больше данных в пределах своих возможностей для их хранения и последующего использования, характер которого предугадать невозможно. Безусловно, в числе таких данных значительное место будут иметь и персональные данные.

В этой связи **существующие законодательные ограничения по обработке персональных данных исключительно в соответствии с первоначально заявленными целями обработки, а также недопустимость объединения различных баз данных с первоначально заявленными и несовместимыми между собою целями обработки, вступает в противоречие с существующей технологией и бизнес-практиками, поскольку оно нивелирует те преимущества, которые предоставляют технологии «Больших данных».** К тому же с учетом современного развития технологий фактическое выполнение данных требований законодательства о персональных данных будет очень сложно проследить.

3.2. «Большие данные» несовместимы с концепцией информированного, конкретного и сознательного согласия как главного основания легитимации обработки персональных данных

Идеальная модель регулирования в сфере законодательства о персональных данных предполагает, что субъект персональных данных имеет возможность самостоятельно принимать решения, касающиеся его информационной сферы, посредством согласия на обработку его персональных данных в отдельных случаях, взвешивая соответствующие риски и выгоды. В связи с этим согласие субъекта персональных данных является главным легитимирующим основанием их обработки[42]. Для того, чтобы согласие субъекта персональных данных могло называться информированным, конкретным и сознательным, необходимо, чтобы ему была предоставлена детальная информация о том, как будут использоваться его персональные данные: цели использования, состав обрабатываемых персональных данных и способы их обработки (ч.4 ст. 9, ч. 7 ст. 14 ФЗ «О персональных данных»).

В эпоху «Больших данных» концепция информированного согласия на обработку персональных данных в значительной степени утрачивает свою эффективность в силу ряда причин: а) невозможности предоставить исчерпывающий объем информации о возможных способах и целях обработки персональных данных; б) невозможности субъекта персональных данных адекватно воспринять такую информацию; в) невозможности индивидуального взаимодействия с огромным множеством организаций, осуществляющих сбор и обработку персональных данных в современном обществе. Рассмотрим данные причины подробнее.

а) Как отмечалось ранее, невозможно заранее предоставить исчерпывающий и конкретный перечень целей, для которых персональные данные могут быть потенциально использованы, в условиях когда «Большие данные» открывают неограниченные возможности к извлечению выгоды от их повторного использования, в том числе путем комбинирования их с иной информацией. Конечно, можно попробовать изложить со-

[42] Article 29 Data Protection Working Party, Opinion 15/2011 on the Definition of Consent, WP 187, July13, 2011, URL: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf (Дата обращения: 14.02.2015); *Bygrave L.* Op.cit. P. 158; *Бачило И.Л.*, *Сергиенко Л.А.*, *Кристальный Б.В.*, *Арешев А.Г.* Указ. соч. С. 24.

ответствующие политики конфиденциальности (политики обработки персональных данных) максимально абстрактным способом, для того, чтобы сохранять пространство для маневра в будущем и по этому пути уже идут многие Интернет-компании. В качестве примера можно привести Политику конфиденциальности Google, в которой указано, что посредством полученных от пользователей данных компания может «поддерживать, защищать, развивать существующие сервисы и создавать новые, а также обеспечивать безопасность Google и наших пользователей. Помимо прочего, эти данные нужны для того, чтобы более точно персонализировать контент, в том числе повышать релевантность результатов поиска и отображаемой рекламы»[43]. Также Google сохраняет за собой право «предоставлять обобщенные обезличенные данные всем пользователям и нашим партнерам, таким как издатели, рекламодатели или связанные сайты. Они могут применяться, например, для того, чтобы проиллюстрировать тенденции использования наших служб»[44]. Очевидно, что из такого описания определить, как будут использоваться персональные данные пользователя, достаточно сложно.

б) Если же начать подробно расписывать все возможные способы использования персональных данных, то соответствующие документы (политики конфиденциальности) окончательно превратятся в нечто, слабо доступное для восприятия. Уже сейчас они состоят из множества страниц, написанных мелким шрифтом, в силу чего сложны для изучения даже профессиональными юристами. Множество исследований показало, что подавляющее большинство пользователей не читают политики конфиденциальности[45]. Большинство из немногих пользователей, кто их читает, не способен понять их содержания и значения[46]. Как отмечается, для того, чтобы понять значительную часть такого рода документов необходимо иметь как минимум высшее образование[47].

Возникает ситуация, получившая в иностранной литературе наименование «Парадокса прозрачности» (*Transparency paradox*), суть которой сводится к тому, что простота и ясность изложения неизбежно сопряжена с упрощениями и утратой важных деталей, а следовательно — с недостатком информации[48]. Взять, например, ситуацию использования технологий «Больших данных» для целей создания и распространения адресной рекламы. Информированное согласие субъекта персональных данных должно предполагать сообщение ему точной информации о видах персональных данных, сбор которых осуществляется; лицах, которым они передается для обработки, условиях такой обра-

---

[43] Политика конфиденциальности Google. Редакция от 31.03.2014 г. URL:http://www.google.com/policies/privacy/ (дата обращения: 14.02.2015).

[44] Там же. Как будет показано далее, тот факт, что соответствующие данные предоставляются в обезличенной форме в эпоху Больших данных не гарантирует невозможности идентификации на их основе конкретного индивида.

[45] *Solove D*. Privacy Self-Management and the Consent Dilemma // Harvard Law Review. No. 126. 2013. P. 1884; *Milne G.*, *Culnan M*. Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices // Journal of Interactive Marketing. No. 18:3. 2004. P. 20-21.

[46] В принципе данная проблема не является новой, она характерна для большинства стандартных договоров, заключаемых потребителями, однако в сфере изучения политик конфиденциальности, которые воспринимаются в ряде зарубежных правопорядков как составная часть договора (Условий использования сервиса) она проявляется наиболее ярко. О проблемах понимания потребителями содержания стандартных условий договоров и сопряженных с этим когнитивных ошибках см. подробнее: *Карапетов А.Г.*, *Савельев А.И.* Свобода договора и ее пределы. Т. 2. М., Статут, 2012.

[47] *Kim N*. Wrap Contracts. Foundations and Ramifications. N.Y.: Oxford University Press, 2013. P. 83.

[48] *Nissenbaum H*. A Contextual Approach to Privacy Online // Daedalus. No. 4. 2011. P. 32-48; *Richards N.*, *King J*. Three Paradoxes of Big Data // Stanford Law Review Online. No. 66. 2013. P. 42-43.

ботки и ее целях; условиях и порядке анонимизации персональных данных (при наличии таковой) и т.д. Очевидно, что времени на изучение такого рода документа в процессе совершения обычной покупки через Интернет-магазин требуется много больше, чем собственно на совершение покупки, а именно возможность сэкономить время является одной из наиболее привлекательных черт электронной коммерции. Получается, что концепция информированного согласия на обработку персональных данных вступает в противоречие с основной ценностью, предоставляемой современными информационными технологиями: оперативностью соответствующих коммуникаций (транзакций).

Наконец, даже если представить, что информация в политике конфиденциальности была поставлена с необходимым уровнем детализации, а также что пользователь прочел и понял содержание документа, оценка им возможных рисков, связанных с соответствующими положениями, сопряжена со значительными сложностями в силу отдаленности и абстрактности возможных негативных последствий. Если негативные последствия курения неплохо поддаются визуализации, в силу чего соответствующие уведомления и изображения больных на упаковках сигарет в некоторых странах обладают немалой эффективностью, визуализировать негативные последствия от обработки отдельных персональных данных с использованием технологий «Больших данных» гораздо труднее[49]. Многие негативные эффекты ненадлежащей обработки персональных данных имеют кумулятивный характер и возникают лишь по прошествии времени, в том числе вследствие комбинирования данных из различных источников. Очень сложно определить, будет ли разглашение (обработка) одних персональных данных, будучи впоследствии объединенной с иной информацией, влечь разглашение чувствительной для субъекта персональных данных информации. Как известно из психологии, способности людей предугадывать, как события отразятся на их благосостоянии в будущем, крайне ограничены[50]. Особенно трудно предугадать, как тот или иной «лайк» в социальной сети, единичный поисковый запрос либо данные GPS об одной поездке могут впоследствии, по прошествии продолжительного времени, отразиться на частной жизни пользователя[51].

Информационная сфера является слишком тонкой материей для того, чтобы рассчитывать на то, что в ней будет разбираться большинство пользователей сети Интернет. Инвестирование своего времени в изучение соответствующих положений политики конфиденциальности является малопродуктивным еще и по той причине, что такие документы подвержены частым изменениям. Нет никаких гарантий, что через некоторое время их положения не изменятся, поскольку практически все политики конфиденциальности содержат оговорки о возможности их изменения в одностороннем порядке[52].

в) Среднестатистический пользователь сети Интернет посещает десятки, а то и сотни веб-сайтов каждый месяц, практически каждый из которых осуществляет сбор и обработку определенной персональной информации о нем. Даже если предположить, что такой пользователь готов в принципе уделять свое время и силы изучению вопросов использования его персональных данных, а соответствующий веб-сайт — предоставлять достоверную и подробную информацию, в итоге все равно получается слишком большой для изучения и понимания объем информации. К тому же существует немало

[49] *Calo R.* Against Notice Skepticism in Privacy (and Elsewhere) // Notre Dame Law Review. No. 87. 2012. P. 1033.

[50] *Solove D.* Op.cit. P. 1891.

[51] *Crawford K., Schultz J.* Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms // Boston College Law Review. No. 55:93, 2014. P.106.

[52] *Kim N.* Op. cit. P. 84.

обработчиков персональных данных, с которыми пользователь может не сталкиваться напрямую, вроде информационных брокеров, рекламных агентств и т.п. Очевидно, что нельзя эффективно управлять своей информационной безопасностью в условиях отсутствия знаний обо всех лицах, которые так или иначе используют персональные данные, а также знаний о том, как именно они это делают. Это лишний раз подтверждает тот факт, что если каждая компания будет скрупулезно описывать все способы и цели обработки персональных данных, современный пользователь — субъект персональных данных не в состоянии этой информацией пользоваться и делать сознательный выбор, который подразумевается Законом о персональных данных. В этой связи показательными являются результаты исследования, проведенного в США: установлено, что среднестатистический американец должен затратить приблизительно 201 час, в стоимостном выражении составляющих в среднем 3 534 долл., на одно только чтение политик конфиденциальности, размещенных на веб-сайтах, которые он посещает. При этом если каждое лицо будет тратить свое время на изучение политик конфиденциальности каждого веб-сайта, которое оно посещает, то общая стоимость потерянного времени в течение года будет составлять порядка 781 млрд. долл.[53].

Все это приводит к неутешительным выводам: **информированное согласие предполагает необходимость принятия субъектом персональных данных ряда дискретных решений на ранних стадиях их обработки (как правило, стадии сбора персональных данных), однако в силу особенностей применения технологий «Больших данных» последствия таких решений невозможно предугадать на данном этапе. В итоге информированное согласие является в современных условиях не более, чем фикцией и не может выполнять роль главного легитимирующего основания для обработки персональных данных.**

3.3. Обезличивание персональных данных не является гарантией их анонимности в эпоху «Больших данных»

Обезличивание персональных данных является одной из мер, направленных на минимизацию рисков причинения вреда гражданам в случае утечки их персональных данных из информационных систем[54]. Обезличивание персональных данных выполняет важную социальную функцию, обеспечивая автономию личности человека и его «недосягаемость» для тех, кто не согласен с высказанными лицом мнениями либо является потенциально враждебным к тем чертам, которые у него присутствуют (определенным заболеваниям, происхождению, убеждениям и т.д.)[55].

Под обезличиванием персональных данных понимаются действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту персональ-

[53] *McDonald A.*, *Cranor L.*, The Cost of Reading Privacy Policies // Journal of Law and Policy for the Information Society. No. 4. 2009. P. 544.

[54] См. пп. «з» п. 1 Постановления Правительства РФ от 21.03.2012 N 211 «Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных» и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами»

[55] Nissenbaum H. The Meaning of Anonimity in an Information Age // The Information Society No. 15:2. 1999. P. 142.

ных данных (ст. 3 Закона о персональных данных). Требования и методы по обезличиванию персональных данных утверждены приказом Роскомнадзора[56]. Среди методов обезличивания данный приказ упоминает следующие методы: введения идентификаторов; изменения состава или семантики; декомпозиции; перемешивания. Конкретный метод выбирается оператором в зависимости от целей и задач обработки персональных данных, учитывая, что обезличивание персональных данных должно обеспечивать не только защиту от несанкционированного использования, но и возможность их обработки, т.е. данные после обезличивания должны обладать рядом свойств, конкретный набор которых зависит от применяемого метода обезличивания. К числу таких свойств относятся помимо всего прочего полнота (сохранение всей информации о персональных данных конкретных субъектов, которая имелась до обезличения) и анонимность (невозможность однозначной идентификации субъектов данных, полученных в результате обезличивания, без применения дополнительной информации). При этом, как указано в приказе Роскомнадзора, одним из свойств применяемого метода обезличивания является обратимость (то есть возможность проведения деобезличивания — приведения данных к исходному виду, позволяющему установить их принадлежность конкретному лицу), а также возможность косвенного деобезличивания (возможность проведения деобезличивания с использованием информации других операторов).

Возникает вопрос, относятся ли деперсонализированные данные к категории персональных данных или представляют собой особый вид данных, на который не распространяется режим персональных данных? Распространенной является точка зрения, согласно которой деперсонализация выводит данные из-под режима персональных и в связи с этим является удобной альтернативой необходимости соблюдения обременительных норм, связанных с обработкой персональных данных[57]. Европейская рабочая группа во вопросам персональных данных пришла к выводу о том, что если анонимизированные данные являются обратимыми, то есть могут быть возвращены к исходному состоянию, то они относятся к категории информации, которая может косвенно определить лицо, а следовательно — являются персональными данными[58].

Российский закон о персональных данных не дает прямого ответа на данный вопрос. Учитывая, что дефиниции персональных данных в европейском и российском праве очень близки, есть основания полагать, что на обезличивания данные распространяется правовой режим персональных данных с некоторой спецификой. Так, Закон о персональных данных предусматривает два специальных правила обезличенных данных: 1) возможность их обработки в статистических и исследовательских целях без согласия пользователя (п. 9 ч. 1 ст. 9) и 2) обезличивание как альтернативу удалению персональных данных по достижении целей обработки (ч. 7 ст. 5)[59].

---

[56] Приказ Роскомнадзора от 05.09.2013 N 996 «Об утверждении требований и методов по обезличиванию персональных данных» (вместе с «Требованиями и методами по обезличиванию персональных данных, обрабатываемых в информационных системах персональных данных, в том числе созданных и функционирующих в рамках реализации федеральных целевых программ») (Зарегистрировано в Минюсте России 10.09.2013 N 29935).

[57] *Carey P.*Data Protection: A Practical Guide to UK and EU Law. Oxford University Press. 2004. P. 58; *Ohm P.* Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization // UCLA Law Review No. 57. 2010. P. 1738.

[58] Opinion 4/2007 on the Concept of Personal Data, WP136 (2007), P. 18. URL:http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf(дата обращения: 14.02.2015).

[59] Законопроектом № 416052-6 о внесении изменений в Федеральный закон № 152 «О персональных данных»предлагается дополнить положением о том, что «Операторы и иные лица, получившие доступ

Учитывая, что большинство методов деперсонализации, указанных в приказе Роскомнадзора, предполагают обратимость, есть основания для вывода о том, что деперсонализация персональных данных не выводит их за рамки действия Закона о персональных данных. В качестве исключения может рассматриваться такой метод деперсонализации как изменение состава или семантики, который производит замену персональных данных результатами статистической обработки, обобщения или удаления части сведений[60].

Так или иначе, на обезличивание (анонимизацию) персональных данных возлагаются большие надежды, поскольку данное средство является, по мнению ряда исследователей, одним из наиболее перспективных способов решения вопроса защите персональных данных в условиях их повсеместной передачи посредством сети Интернет[61]. Однако в эпоху «Больших данных», когда становится возможным идентифицировать личность посредством установления корреляций между несколькими фрагментами данных, эффективность данного способа вызывает сомнения[62]. При этом не важно, какой именно метод анонимизации данных использован. Любой идентификатор или любая информация об относительно уникальном качестве лица (например, его музыкальных предпочтениях или посещенных местах) может служить основанием для «опознания» данного лица в различных базах данных. Риски деанонимизации в значительной степени увеличились в связи с появлением социальных сетей и иных веб-сайтов, где люди оставляют значительное количество информации о себе. Впрочем, практически любое действие пользователя в сети Интернет может служить в качестве связующего звена к идентификации его личности, поскольку оно оставляет так называемый цифровой след (*digital finger print* — букв. «цифровые отпечатки пальцев»)

Например, в свое время компания AOL сделала общедоступными совокупность старых поисковых запросов с намерением дать возможность их использования в исследовательской деятельности. Набор данных из 20 миллионов поисковых запросов 650 тысяч пользователей за период с 1 марта по 31 мая 2006 г. был тщательно анонимизирован: личные данные пользователей в виде имен и IP адресов были удалены и замещены уникальными цифровыми идентификаторами. Однако на основании сопоставления различных запросов удалось установить личности ряда пользователей. Можно привести и иной пример. Известный Интернет-сервис проката фильмов Netflix выпустил 100 миллионов записей о прокате от полумиллиона пользователей, личные идентификаторы которых были удалены, с целью проведения конкурса на улучшение системы рекомендаций фильмов. Однако сравнив данные Netflix с иными общедоступными источниками (в частности с данными об оценках пользователями фильмов на известном веб-сайте IMDb), исследователи пришли к выводу, что на основе всего шести оценок фильмов можно установить личность пользователя в 84% случаев, а зная дату оцен-

---

к персональным данным не обязаны обеспечивать их конфиденциальность в случае если они были обезличены» (ч. 2 ст. 7) URL: http://asozd.duma.gov.ru/main.nsf/(Spravka)?OpenAgent&RN=416052-6 (Дата обращения: 14.02.2015).

[60] *Жаринов Р., Трифонова Ю.* Возможности обезличивания персональных данных в системах, использующих реляционные базы данных // Управление, вычислительная техника и информатика. № 2 (32), 2014. P. 189.

[61] Cloud Computing Law / ed. by Millard C. N.Y.: Oxford University Press, 2013. P. 169-178.

[62] *Narayanan A., Shmatikov V.* Robust De-Anonymization of Large Sparse Datasets // The University of Texas. 2008 IEEE Symposium on Security and Privacy. URL:https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf(дата обращения: 14.02.2015); *Dwork C.* Differential Privacy. ICALP '06. Proceedings of the 33rd International Conference on Automata, Languages and Programming. Berlin: Springer, 2006.

ки — с 99% точности[63]. Особенно подвержены деанонимизации люди, пользующиеся социальными сетями за счет возможности проследить так называемый «социальный граф» — дружеские связи между пользователями таких сетей: исследования доказали возможность идентификации анонимных пользователей социальных сетей исключительно на основании анализа социального графа[64].

Таким образом, два основных фактора: рост производительности и доступности вычислительных мощностей, а также огромный массив доступной в сети Интернет личной информации — обусловливают техническую возможность деанонимизации даже тщательно обезличенных данных, ибо любые обезличенные данные всегда имеют какой-либо атрибут, относимый к личности. Как отмечается, в эпоху «Больших данных» данные могут быть либо представлять ценность для обработки, либо быть анонимными, но одновременно и тем, и другим — никогда[65]. Чем больше степень обезличивания данных, тем меньше ценность таких данных для анализа. Если информация в современном мире действительно представляет собою новую «нефть», то наивно ожидать, что компании будут ее уничтожать вместо того, чтобы извлекать из нее выгоду.

Однако реальная проблема кроется еще глубже: в подавляющем большинстве случаев Интернет-компаниям или иным лицам, заинтересованным в получении сведений, отражающих признаки лица, не требуется знать его имя. Как отмечается, «если компания имеет порядка 100 единиц информации обо мне, которые оказывают влияние на то, как она строит свои отношения со мною в цифровой среде, какая разница, знают они мое имя или нет?»[66]. В современных технических реалиях компании не обязательно знать имя лица, чтобы персонализировать свое отношение к нему и предлагать соответствующие товары (услуги). Реальная «оффлайн» личность лица не имеет особого значения в сети Интернет, имеют значение те характеристики личности, в которых проявляется поведение и предпочтения лица в сети. «Большие данные» позволяют создавать детальные портреты людей, без необходимости раскрывать при этом их реальные личности. В той мере, в какой эти данные учитываются при принятии решений в отношении такой личности (например, принятие решения о заключении или отказе в заключении договора, определение индивидуальной стоимости товара для него, направлении персонализированной рекламы или иного контента для него и т.п.), защита персональных данных, обеспечиваемая посредством их обезличивания, мало что значит. В иностранной литературе по этой причине уже высказываются мнения о том, что необходимо переходить от регулирования собственно персональных данных к регулированию оборота информации в целом[67]. Так или иначе, **в новых технологических реалиях обезличивание данных уже не может выполнять функцию эффективного средства защиты персональных данных и в более глобальном смысле — частной жизни граждан.**

Сказанное не означает, что обезличивание персональных данных является бесполезным и от него следует отказаться, речь идет о том, что оно не должно рассматриваться в качестве средства, безусловно достаточного для эффективной защиты персональных данных в эпоху «Больших данных», и возлагать на него чрезмерные надежды.

---

[63] *Майер-Шенбергер В., Кукьер К.* Указ. соч. С. 162.

[64] *Narayanan A., Shmatikov V.*, De-Anonymizing Social Networks. 30th IEEE Symposium on Security & Privacy, 2009. URL: http://userweb.cs.utexas.edu/-shmat/shmat-oak09.pdf(Дата обращения: 14.02.2015).

[65] *Ohm P.* Op.cit. P. 1704.

[66] *Lane J., Stodden V., Bender S., Nissenbaum H.* Op.cit. P. 70.

[67] *Gutwirth S., Hert P.* Regulating Profiling in Democratic Constitutional State in Profiling the European Citizen: Cross-Disciplinary Perspectives / ed. by Hildebrandt M. Dordrecht: Springer, 2008. P. 289.

Выводы

Технологии «Больших данных» ознаменовали момент, когда понятие «информационное общество» приобрело полноценный смысл. Информация приобрела статус ценного актива — своего рода новой нефти, — выступающей движущей силой информационного общества подобно тому, как традиционная нефть выступала главным ресурсом в эпоху индустриального общества. Технологии «Больших данных» сулят большие выгоды в самых различных сферах: появление новых бизнес-моделей, построенных на индивидуальном отношении к клиенту; совершенствование системы здравоохранения, улучшение криминогенной ситуации в крупных городах, борьба с мошенническими действиями и т.д. Однако «Большие данные» имеют и темную сторону, обладая значительным потенциалом для вторжения в частную жизнь граждан. Проблемы, связанные с влиянием технологий «Больших данных» на применение законодательства о персональных данных, признаны в Европе. Европейская рабочая группа по вопросам персональных данных указала, что вызовы, бросаемые технологиями «Больших данных», требуют инновационного подхода к толкованию и применению базовых принципов законодательства о персональных данных, а также их дальнейшего совершенствования, хотя, по ее мнению, на данном этапе рано говорить о том, что данные принципы абсолютно не действуют в новых реалиях[68].

В настоящей статье было продемонстрировано конфликтное состояние технологий Больших данных с законодательством о персональных данных на примере трех положений последнего: принципа минимизации данных и ограничения обработки заранее определенной целью; концепции информированного согласия как ключевого основания обработки, а также возможности обезличивания персональных данных с целью исключения полученных данных из-под действия законодательства о персональных данных. Однако было бы наивным полагать, что этим проблемы исчерпываются. На самом деле технологии «Больших данных» обнажают очевидный факт: законодательство о персональных данных в том виде, в каком оно было сформулировано еще в Конвенции 1981 г., становится все менее и менее адекватным современным технологическим реалиям и нуждается в существенной переработке. Косметическое или точечное изменение существующего регулирования в сфере защиты персональных данных не способно сделать его эффективным, а способно лишь увеличить степень его отрыва от реальности.

Переосмыслению должны подвергнуться такие базовые категории, как понятие персональных данных и понятие оператора персональных данных. В условиях, когда сбор сведений о пользователях носит массовый характер, даже самый безобидный фрагмент такой информации (сведения о посещении сайта или совершении покупки) будучи соединенным с другой подобной информацией, способен дать гораздо больше сведений о лице, чем совокупность его анкетных данных. Достойны ли такие единицы информации особого регулирования путем придания им статуса персональных данных? Либо же имеет смысл выделить их в особую категорию с отдельным регулированием? Влияет ли анонимность или использование псевдонима на возможность квалификации соответствующей информации как персональных данных? Эти вопросы требуют разрешения.

[68] Statement on Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU. Article 29 Data Protection Working Party, WP221. September 16, 2014. URL: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf (датаобращения: 14.02.2015).

Беспрецедентные масштабы циркулирующей в цифровой форме информации об индивидах повлекли появление за рубежом новых игроков на рынке информации: информационных брокеров, которые на основе стекающейся к ним из различных Интернет-сервисов данных составляют детальные профайлы граждан и предоставляют доступ к ним заинтересованным лицам. Существующее регулирование, построенное на понятии оператора персональных данных, не учитывает существенных различий между 1) лицами, осуществляющими первичный сбор данных (различного рода Интернет-сервисами); 2) лицами, агрегирующими на профессиональной основе такие данные посредством технологий «Больших данных» в профайлы и 3) лицами, приобретающими доступ к таким профайлам для собственных нужд (например, финансовыми учреждениями для оценки платежеспособности клиента). Все указанные лица являются операторами персональных данных, однако их действия имеют различный характер и могут влечь различные последствия для субъекта персональных данных с точки зрения возможного вреда от их ненадлежащей обработки.

Указанные вопросы нельзя решить, не ответив первоначально на главный вопрос: являются ли персональные данные товаром или они являются неотчуждаемым неимущественным благом? В пользу каждой трактовки можно привести множество аргументов. Однако одним из наиболее значимых является то, что бóльшая часть успешных бизнес-моделей в сети Интернет основана на использовании персональных данных в качестве «валюты», которой пользователь расплачивается за возможность использования соответствующего сервиса. Именно данные пользователей являются одним источником многомиллиардных доходов социальных сетей, поисковых сервисов и иных IT-компаний. Однако этот факт игнорируется большинством пользователей, как, впрочем, и российским законодательством, не признающим предоставление персональных данных на обработку в качестве встречного предоставления для целей квалификации договора в качестве возмездного (ст. 424 ГК РФ). Так или иначе, от ответа на данный вопрос будет во многом зависеть решение и всех остальных, ранее обозначенных проблем. Попытки поиска ответов на поставленные проблемы будут предприняты в дальнейших работах автора.

Библиография

Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. М.: Манн, Иванов и Фербер, 2014. 240 с.

Бачило И.Л., Сергиенко Л.А., Кристальный Б.В., Арешев А.Г. Персональные данные в структуре информационных ресурсов. Основы правового регулирования. Минск: Беллитфонд, 2006. 474 с.

Брагинский М.И., Витрянский В.В. Договорное право. Т.3: Договоры о выполнении работ и оказании услуг. М.: Статут, 2011. 872 с.

Войниканис Е. Право интеллектуальной собственности в цифровую эпоху. Парадигма баланса и гибкости. М.: Юриспруденция, 2013. 552 с.

Жаринов Р., Трифонова Ю. Возможности обезличивания персональных данных в системах, использующих реляционные базы данных // Управление, вычислительная техника и информатика. № 2 (32), 2014. С. 188–194.

Bygrave L. Data Privacy Law: An International Perspective. Oxford University Press, 2014. 233 p.

Calo R. Against Notice Skepticism in Privacy (and Elsewhere) // Notre Dame Law Review. No. 87. Pp. 1027-1072.

Carey P. Data Protection: A Practical Guide to UK and EU Law. Oxford University Press, 2004. 532 p.

Crawford K., Schultz J. Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms // Boston College Law Review. No. 55:93, 2014. Pp. 93-128.

Gutwirth S., Hert P. Regulating Profiling in Democratic Constitutional State in Profiling the European Citizen: Cross-Disciplinary Perspectives / ed. by M. Hildebrandt, S. Gutwirth. Dordrecht, Springer, 2008. 374 p.

Joh E. Policing by Numbers: Big Data and the Fourth Amendment // Washington Law Review. 2014. No. 89:35. Pp. 35–68.

Kim N. Wrap Contracts. Foundations and Ramifications. N.Y., Oxford University Press, 2013. 240 p.

Lane J., Stodden V., Bender S., Nissenbaum H. Privacy, Big Data and the Public Good: Frameworks for Engagement. Cambridge University Press, 2014. 344 p.

Milne G., Culnan M. Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices // Journal of Interactive Marketing. 2004. no. 18:3. P. 15–29.

Millard C. (ed.). Cloud Computing Law. Oxford University Press, 2013. 416 p.

McDonald A., Cranor L., The Cost of Reading Privacy Policies // Journal of Law and Policy for the Information Society. 2009. No. 4. P. 543–568.

Narayanan A., Shmatikov V. Robust De-Anonymization of Large Sparse Datasets // The University of Texas. IEEE Symposium on Security and Privacy, 2008. P. 111–125.

Narayanan A., Shmatikov V., De-Anonymizing Social Networks. IEEE Symposium on Security and Privacy, 2009. P. 173–187.

Nissenbaum H. The Meaning of Anonimity in an Information Age // The Information Society. 1999. No. 15:2. P. 141–144.

Nissenbaum H. A Contextual Approach to Privacy Online // Daedalus, 2011. No. 4. P. 32–48.

Ohm P. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization // UCLA Law Review No. 57. 2010. P. 1701–1777.

Rajesh D. The NSA and Accountability in an Era of Big Data // Journal of National Security Law & Policy, 2014, no. 7. P. 301–310.

Richards N., King J. Three Paradoxes of Big Data // Stanford Law Review Online. No. 66. 2013. P. 41–46.

Solove D. Privacy Self-Management and the Consent Dilemma // Harvard Law Review. No. 126. 2013. P. 1880–1903.

The Issues of Implementing Legislation on Personal Data in the Era of Big Data

Alexander Savelyev

Senior Legal Researcher, Center on Information Law, National Research University Higher School of Economics, Legal Advisor of IBM Russia/CIS, Member of Advisory Board at the Federal Services for Supervision of Communications, Information Technology and Mass Media, Candidate of Juridical Sciences. Address: Myasnitskaya Str. 20, Moscow, 101000, Russian Federation. E-mail: garantus@rambler.ru.

Abstract

This paper analyses the impact of Big Data technologies on data protection legislation, which represents one of the main legal outposts of privacy in digital environment. The paper describes the origin of what is currently called Big Data, its definition and examples of its application in various spheres. The main focus of the paper is on the analysis of compatibility of Big Data technologies with the core principles of data protection legislation. Based on the analysis, the author comes to a conclusion that such principles as purpose limitation, data minimization; informed consent as a main basis for processing personal data are substantially eroded by Big Data. Purpose limitation and data minimization are at odds with the concept of data reuse, which underpins the philosophy of the Big Data age. Informed consent is impossible in the situations where a specific goal of data processing cannot be provided, and it cannot be provided due to unpredictable nature of potential data uses of in Big Data age: to limit processing

data by specific purposes means rejection of the benefits of Big Data. In addition, the point is made that contrary to the popular view, anonimity of personal data is not an effective solution of existing problems with personal data in Big data era, due to widely available opportunities for de-anonymization provided by cheap computing power and a vast amount of information currently available on the Internet. The paper has a purpose of highlighting the problems in data protection legislation and initiating discussions. It is expected that possible solutions to them will be a subject of subsequent papers.

Keywords

personal data, Big Data, anonymization, purpose limitation, profiling, data brokers, data subject consent, data minimization

Citation: Savelyev A.I. (2015) The Issues of Implementing Legislation on Personal Data in the Era of Big Data. *Pravo. Zhurnal Vysshey shkoly ekonomiki*, no.1, pp. 43–66 (in Russian)

JEL: K39

References

Bachilo I.L., Sergienko L.A., Kristal'nyy B.V., Areshev A.G. (2006) *Personal'nye dannye v strukture informatsionnykh resursov. Osnovy pravovogo regulirovaniya* [Personal Data in Information Resource. Fundamentals of Legal Regulation]. Minsk, Belitfond, 474 p.

Braginskiy M.I., Vitryanskiy V.V. (2011) *Dogovornoe pravo. Tom 3: Dogovory o vypolnenii rabot i okazanii uslug* [Contract Law. Vol. 3: Work and Service Contracts]. Moscow, Statut, 872 p.

Bygrave L. (2014) *Data Privacy Law: An International Perspective*, Oxford University Press, 233 p.

Carey P. (2004) Data *Protection: A Practical Guide to UK and EU Law*. Oxford University Press. 532 p.

Calo R. Against Notice Skepticism in Privacy (and Elsewhere). *Notre Dame Law Review*, no. 87, pp. 1027–1072.

Crawford K., Schultz J. (2014) Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review*, no. 55, p. 93–128.

Gutwirth S., Hert P. (2008) Regulating Profiling in Democratic Constitutional State. Hildebrandt M., Gutwirth S. (eds.) *Profiling the European Citizen: Cross-Disciplinary Perspectives*. Dordrecht, Springer. 374 p.

Joh E. (2014) Policing by Numbers: Big Data and the Fourth Amendment. *Washington Law Review*, no. 89, p. 35–68.

Kim N. (2013) *Wrap Contracts. Foundations and Ramifications*. Oxford University Press, 240 p.

Lane J., Stodden V., Bender S., Nissenbaum H. (2014) *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. Cambridge University Press, 344 p.

Mayer-Schönberger V., Cukie K. (2014) *Bol'shie dannye. Revolyutsiya, kotoraya izmenit to, kak my zhivem, rabotaem i myslim* [Big Data: A Revolution that Will Transform How We Live, Work and Think]. Moscow: Mann, Ivanov e Ferber, 240 p.

McDonald A., Cranor L. (2009) The Cost of Reading Privacy Policies. *Journal of Law and Policy for the Information Society*, no. 4, pp. 543–568.

Milne G., Culnan M. (2004) Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices. *Journal of Interactive Marketing*, no. 18, pp. 15–29.

Millard C. (ed.) (2013) *Cloud Computing Law*. Oxford University Press, 416 p.

Narayanan A., Shmatikov V. (2008) Robust De-Anonymization of Large Sparse Datasets. *The University of Texas. IEEE Symposium on Security and Privacy*, pp. 111–125.

Narayanan A., Shmatikov V. (2009) *De-Anonymizing Social Networks. University of Texasx IEEE Symposium on Security and Privacy*, pp. 173–187.

Nissenbaum H. (1999) The Meaning of Anonimity in an Information Age. *The Information Society*, no. 15, pp. 141–144.

Nissenbaum H.A. (2011) Contextual Approach to Privacy Online. *Daedalus*, no. 4, pp. 32–48.

Ohm P. (2010) Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, no. 57, pp. 1701–1777.

Rajesh D. (2014) The NSA and Accountability in an Era of Big Data. *Journal of National Security Law and Policy*, no. 7, pp. 301–310.

Richards N., King J. (2013) Three Paradoxes of Big Data. *Stanford Law Review Online*, no. 66, pp. 41–46.

Solove D. (2013) Privacy Self-Management and the Consent Dilemma. *Harvard Law Review*, no. 126, pp. 1880–1903.

Voynikanis E. (2013) *Pravo intellektual'noy sobstvennosti v tsifrovuyu epokhu. Paradigma balansa i gibkosti* [Copyright in Digital Era]. Moscow, Jurisprudentsia, 552 p.

Zharinov R., Trifonova Yu. (2014) Vozmozhnosti obezlichivaniya personal'nykh dannykh v sistemakh, ispol'zuyushchikh relyatsionnye bazy dannykh [Anonymisation of Personal Data in the Systems Applying Relational Databases]. *Upravlenie, vychislitel'naya tekhnika i informatika*, no. 2 (32), pp. 188–194.