М.А. Рожкова, В.Н. Глонина

ПЕРСОНАЛЬНЫЕ И НЕПЕРСОНАЛЬНЫЕ ДАННЫЕ В СОСТАВЕ БОЛЬШИХ ДАННЫХ

Аннотация. *В данной статье авторы рассматривают большие данные в контексте правового режима различных категорий данных, входящих в их состав. Авторы анализируют понятия «большие данные», «персональные данные», «неперсональные данные», «анонимизированные данные», их характеристики, основные подходы к их пониманию, а также основные проблемы, возникающие на практике в связи с использованием больших данных. Проведя анализ зарубежной и национальной доктрины, законодательства и судебной практики, авторы приходят к выводу, что ценность для использования в составе больших данных имеют как персональные, так и неперсональные данные, в то же время требуется адаптировать законодательство о персональных данных для целей аналитики больших данных, в частности ввести механизм анонимизации данных.*

Ключевые слова: *большие данные, персональные данные, неперсональные данные, анонимизированные данные.*

Большие данные (далее – *big data*) в современном мире рассматриваются как один из ценнейших активов и серьезное конкурентное преимущество компаний. Использование big data позволяет компаниям внедрять новые бизнесы-процессы и ускорять имеющиеся, разрабатывать новые продукты и услуги, эффективно персонализировать сервисы, повышать качество принятия управленческих решений. Такие преимущества оказывают прямое влияние на успех бизнеса, работающего с *big data*. Показательными примерами эффективного использования *big data* являются такие *IT*-гиганты, как *Facebook*, *Google*, *Apple*, *Amazon* и т.д.

Вместе с тем с правовой точки зрения использование *big data* в настоящее время порождает множество рисков, вытекающих из неоднозначного правового режима, а также отсутствия единства в понимании сущности *big data* и их соотношения с уже имеющимися в законодательстве категориями – «информация», «персональные данные», «база данных», «ноу-хау», «тайна» (коммерческая, врачебная, банковская и т.д.).

Ярким примером возникающих в практике проблем является тянущееся с начала 2017 г. дело «*ВКонтакте против Double Data*»[1], в котором оспаривается правомерность действий стартапа *Double Data*, осуществляющего сбор данных, размещенных пользователями в социальной сети «ВКонтакте» (скрапинг)[2]. Суды разных инстанций по-разному высказались относительно допустимости скрапинга и последующего использования данных, размещаемых пользователями в социальной сети. При этом несмотря на то, что данное дело рассматривается в контексте законодательства об интеллектуальной собственности (в аспекте нарушения смежных прав на базу данных), дело затрагивает вопросы, важные с точки зрения защиты персональных данных и права граждан на частную жизнь.

Существуют и примеры привлечения к ответственности за нарушение требований законодательства о персональных данных компаний, использующих *big data*. Так, в деле *Роскомнадзор против НБКИ*[3] суды установили нарушение законодательства о персональных данных в деятельности скорингового агентства, использовавшего в рамках своих *big data* проектов данные пользователей из социальных сетей.

Одновременно с этим в последнее время наблюдается множество разнообразных законотворческих инициатив, направленных на регулирование отношений в сфере *big data* (законопроект «О больших пользовательских данных»[4] (далее – Законопроект о БД), законопроекты ФРИИ[5] и «Сколково»[6]). Но подобные предложения нередко формируются без четкого понимания того, что собой представляют *big data*.

[1] Дело № А40-18827/2017 (http://kad.arbitr.ru/Card/1f33e071-4a16-4bf9-ab17-4df80f6c1556).

[2] См. об этом: *Рожкова М.А.* Скрапинг/парсинг интернет-ресурсов – что это такое и законно ли это? [Электронный ресурс] // Закон.ру. 2019. 30 дек. (URL: https://zakon.ru/blog/2019/12/30/skraping__parsing_internet-resursov__chto_eto_takoe_i_zakonno_li_eto).

[3] Дело № А40-5250/2017 (http://kad.arbitr.ru/Card/eb1907d9-be95-4b0e-85c7-0481aef89b31).

[4] Законопроект № 571124-7 «О внесении изменений в Федеральный закон «Об информации, информационных технологиях и о защите информации»», внесенный в Государственную Думу 23.10.2018 (http://sozd.duma.gov.ru/bill/571124-7).

[5] Проект федерального закона о внесении изменений в Федеральный закон «Об информации, информационных технологиях и о защите информации и отдельные законодательные акты Росиийской Федерации» (https://www.iidf.ru/media/articles/fond/frii-proekt-zakona-o-regulirovanii-personalnykh-dannykh/).

[6] Проект федерального закона о внесении изменений в федеральные законы «Об информации, информационных технологиях и о защите информации» и «О персональных данных» (https://sk.ru/foundation/legal/p/03.aspx).

Многие юристы рассматривают *big data* исключительно как огромные массивы разнообразной информации (данных), чему способствовало растиражированное во многих публикациях употребление Клиффордом Линчем[1] этого термина применительно к взрывному росту мировых объемов информации и многообразию данных. Между тем огромный физический объем и разнообразие типов данных представляют собой только некоторые – наиболее известные характеристики такого многоаспектного явления, как *big data*. И для того, чтобы дать ему правовую характеристику, необходимо изучить и другие его аспекты.

Big data: подходы к пониманию и основные характеристики

Один из наиболее явных проблемных моментов в понимании *big data* – это употребление данного термина в нескольких значениях.

Как указывалось выше, традиционно *big data* рассматривают в качестве огромного массива информации, включающего в свой состав самые разнообразные данные. Причем выделение *big data* в качестве *самостоятельной категории данных* весьма условно: по большому счету в их состав могут попадать абсолютно все виды данных. При этом часть информации может включаться в виде структурированных, т.е. упорядоченных определенным образом, данных (например, они могут быть собраны в таблицы, что позволяет применять в отношении них визуальный или автоматизированный анализ). Но большая часть информации поступает в виде данных неструктурированных (англ. *unstructured data*), к которым относят, в частности, данные из соцсетей, видео- и аудиофайлы, данные *GPS*, спутниковые изображения, данные о перемещении мобильного абонента, данные с серверов, файлы *PDF* и проч. В отношении этих, а также полуструктурированных данных затруднено использование программ, предназначенных для работы со структурированными данными, – требуется разработка специальных технических решений.

Последнее обстоятельство отчасти и стало причиной того, что сегодня все большее распространение получает понимание *big data* как цифровых технологий[2], в частности новых технологий сбора, накопле-

[1] Редактор журнала «Nature», которому приписывается введение этого термина в обиход в 2008 г. (см.: https://www.nature.com/articles/455001a).

[2] Это нашло отражение в федеральном проекте «Цифровые технологии» национальной программы «Цифровая экономика Российской Федерации», в рамках которо-

ния, хранения, аналитики данных[1]. Причем особая важность усматривается в большей степени за *аналитикой больших данных* (далее – *big data analytics*), которая предполагает действия по структуризации данных, созданию алгоритмов анализа данных, агрегации и анализу данных, выявлению связей между данными, установлению закономерностей и скрытых тенденций, построению прогнозов и т.п., а в публикациях именуется «развивающейся (новой) формой производства знаний»[2]. Именно *big data analytics* делает большие данные экономическим активом: огромные объемы необработанной информации («сырые» данные), требующие существенных затрат на хранение, обладают только потенциальной коммерческой значимостью, тогда как реальную коммерческую ценность *big data* приобретают после их обработки, включая аналитику, которая позволяет использовать большие данные для решения той или иной научной, социальной либо коммерческой задачи.

В качестве примера можно указать на разработку российского стартапа *Synqera*[3], созданного на основе исследования, показавшего, что покупатели тратят миллионы на импульсные покупки. Разработанная стартапом вычислительная платформа за 40 секунд – время ожидания, которые покупатель в среднем проводит на кассе в ожидании оплаты товаров – анализирует информацию о каждом покупателе: историю его покупок, покупательские предпочтения, возраст, пол и даже настроение (на кассах магазинов сенсорные экраны с датчиками распознают эмоции покупателей). Полученный результат обогащается бизнес-информацией об акциях или скидках магазина, а также данными из открытых источников (из соцсетей или о погоде). По результатам каждому покупателю отправляются таргетированные сообщения, предоставляются персональные скидки и специальные предложения и т.п., способствующие совершению этим покупателем упомянутых импульсных покупок. Таким образом, вычислительная платформа позволяет использовать время, проведенное покупателем на кассе, для стимуляции таких покупок.

го в перечне сквозных цифровых технологий упоминаются технологии больших данных (https://digital.gov.ru/ru/activity/directions/878/#section-description).

[1] См. подробнее: *Рожкова М.А.* Характеристики больших данных, значимые для целей гражданского права // Хозяйство и право. 2019. № 6. С. 21–28.

[2] См., например: *Mayer-Schönberger V. and Cukier K.* Big Data: A Revolution that Will Transform how We Live, Work, and Think. Houghton Mifflin Harcourt, 2013.

[3] *Synqera* представила решение для роста покупательской лояльности (https://www.retail-loyalty.org/news/synqera-predstavila-vysokotekhnologichnoe-reshenie-dlya-rosta-pokupatelskoy-loyalnosti-/).

Отдельно стоит обратить внимание на то, что применительно к *big data analytics* приоритетное значение приобретает не количественная, а качественная сторона данных. Это проявляется, в частности, в том, что для решения различных задач подразумевается задействование разных доступных для аналитики данных – собранных в различном контексте и полученных из многих источников (например, современные автомобили сегодня способны накапливать данные о водителе, погоде и окружающей среде, самом авто и допущенных системой ошибках, подключенных устройствах и др.). При этом *big data analytics* допускает использование одних и тех же доступных для аналитики данных для достижения различных целей, для этого они соответствующим образом трансформируются и агрегируются – с добавлением новых наборов данных или без таковых. То есть анализу могут быть подвергнуты все доступные конкретной компании данные либо выборки из собранных и накопленных данных – в зависимости от поставленной задачи. Это позволяет акцентировать внимание на том, что упомянутые данные могут использоваться бесконечное количество раз и в целях, которые заранее сложно предвидеть.

Примечательно, что проведенные ранее исследования[1] позволили разграничивать данные на две группы – «данные в динамике» (англ. *data in motion*) и «данные в статике» (англ. *data in statics*). Различия между этими группами данных усматриваются в следующем: 1) «данные в статике» могут быть получены из архивов, резервных копий и т.д., в то время как «данные в динамике» собираются в режиме реального времени с помощью технических датчиков, автоматизированных процессов или механического ввода; «данные в динамике» могут сохраняться или не сохраняться, изменяться, повреждаться, раскрываться и т.д.; 2) «данные в статике» хранятся локально, преобразования проводятся локально, когда «данные в динамике» активно перемещаются, агрегируются, для преобразования таких данных используются разнообразные технологии, задействованы различные участники; 3) процессы анализа «данных в статике» не основаны на изменениях и обновлениях, осуществляющихся в режиме реального времени, в то время как «данные в динамике» могут зависеть от контекста, поступать от технических датчиков, а также технологий, позволяющих осуществлять потоковую передачу данных в режиме реального времени. При этом для *big data analytics* могут использоваться и те и другие разновид-

[1] Informed consent and data in motion. Accenture. 2016 (https://www.accenture.com/_acnmedia/PDF-30/Accenture-Informed-Consent-Data-Motion.pdf).

ности названных данных со всеми присущими им характеристиками. Так, в качестве разновидности «данных в статике» можно привести, например, широту и долготу городов – они не меняются, но могут публиковаться пользователями в социальных сетях и таким образом попадать в число больших данных.

Таким образом, *big data* можно рассматривать как постоянный поток огромных объемов информации, непрерывно поступающий из различных источников. Но нельзя забывать и о другой трактовке этого понятия, которая становится сегодня основной, – понимании *big data* как технологий сбора, накопления, обработки и аналитики данных, представляющих собой нередко альтернативу существующим нынче способам и методам.

Источники и виды данных в составе *big data*

Говоря об источниках *big data*, их можно условно объединить в две основные группы: технические и социальные.

Технические источники создают порядка 90% всей новой информации. Эта группа охватывает, в частности, *интернет вещей* (англ. *Internet of Things, IoT*), включая *промышленный интернет* (англ. *Industrial Internet of Things, IIoT*)), который «поставляет» информацию со всевозможных действующих датчиков, контроллеров, приборов учета потребления, устройств, устройств аудио- и видеорегистрации, измерительных комплексов и проч., *искусственный интеллект* (англ. *Artificial Intelligence, AI*) и *машинное обучение* (англ. *Machine Learning, ML*);

Социальные источники, включающие, в частности, *социальные медиа* (англ. *social media*), охватывают разнообразные способы электронной коммуникации: социальные сети, виртуальные миры, специализированные форумы, профессиональные соцсети, блоги, фотохостинги, сайты отзывов, сайты знакомств и проч. (информация образуется из потока постов, комментариев, лайков, поисковых запросов, оценок, фото, аудио- и видеозаписей, отзывов и проч.); *розничную торговлю* (англ. *retail*), предоставляющую информацию о совершенных транзакциях, сведения из товарных чеков, из дисконтных карт и карт лояльности покупателей, из *RFID*-меток и проч.; *здравоохранение*, собирающее сведения о поставленных диагнозах и предложенных методиках лечения, восприимчивости пациентов к медицинским препаратам, об оценке эффективности этих препаратов и проч., что находит отражение в медицинских картах, результатах лабораторных исследований и т.д.

Основываясь на критерии источника, данные в составе *big data* можно разделить на две категории – данные, *поступающие из технических источников*, и данные, *поступающие из социальных источников*.

Получила широкое распространение градация данных на *государственные* и *негосударственные* (*частные*).

Под *государственными данными* подразумевается не информация ограниченного доступа (например, государственная тайна), а разного рода сведения, связанные с деятельностью государственных органов, а также информация, созданная или собранная этими органами в пределах своих полномочий. Уже сейчас в отношении информации, генерируемой в процессе деятельности госорганов, установлен режим открытых данных[1]. Новейшие законодательные акты призваны закрепить соответствие правового режима государственных данных принципу *open by default*, который предполагает открытость данных «по умолчанию», если в отношении данных прямо не введен иной режим (например, режим государственной тайны)[2]. В частности, Концепция создания и функционирования национальной системы управления данными (далее – НСУД) и разрабатываемый на ее основе законопроект прямо предусматривают, что государственные данные[3] будут размещаться в формате открытых данных и станут доступны для обработки третьими лицами.

Применительно к *негосударственным данным* специального регулирования не предусмотрено, но нужно учитывать, что негосударственные данные тоже могут существовать в форме открытых данных. Так,

[1] См. Федеральный закон «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления», Концепцию открытости федеральных органов исполнительной власти (распоряжение Правительства РФ от 30.01.2014 № 93-р), перечень общедоступной информации, обязательно раскрываемой в форме открытых данных (распоряжение Правительства РФ от 10.07.2013 № 1187-р).

[2] См. План Министерства финансов Российской Федерации по реализации мероприятий в области открытых данных в 2018–2019 гг. (утв. Минфином России 28.12.2018), постановление Правительства РФ от 03.06.2019 № 710 «О проведении эксперимента по повышению качества и связанности данных, содержащихся в государственных информационных ресурсах», распоряжение Правительства РФ от 03.06.2019 № 1189-р «Об утверждении Концепции создания и функционирования национальной системы управления данными и плана мероприятий («дорожной карты») по созданию национальной системы управления данными на 2019–2021 гг.».

[3] В Концепции они определены как «информация, содержащаяся в информационных ресурсах органов и организаций государственного сектора, а также в информационных ресурсах, созданных в целях реализации полномочий органов и организаций государственного сектора».

согласно п. 4 ст. 7 Закона об информации к *открытым данным* относится информация из любых источников, если соблюдаются следующие условия: 1) информация размещается в Интернете ее обладателем; 2) имеется возможность осуществлять автоматизированную обработку информации без предварительных изменений человеком в целях повторного ее использования; 3) размещение информации не должно нарушать требования законодательства (в том числе законодательство о персональных данных).

Примечательно, что высказываются различные позиции в отношении возможности использования информации, размещенной в Интернете *в форме открытых данных*.

С одной стороны, как уже было отмечено, для использования государственных открытых данных, как правило, прямо предусматривается возможность их использования свободно, бессрочно, безвозмездно и без ограничения территории использования, причем как в некоммерческих, так и в коммерческих целях. Такой подход закреплен, например, в отношении открытых данных Роскомнадзора (они могут быть использованы с некоторыми ограничениями – только в законных целях, без искажения при использовании, при сохранении ссылки на источник информации[1]). Обращают на себя внимание открытые данные Роспатента: ведомство гарантирует, что эти данные не являются интеллектуальной собственностью третьей стороны, предоставляя их на основе лицензии (*Creative Common Attribution 3.0*), т.е., по сути, рассматривая открытые данные в качестве объектов интеллектуальной собственности, использование которых требует разрешения правообладателя.

С другой стороны, в отношении открытых данных, которые не относятся к государственным, ситуация иная, и, например, использование личных данных, опубликованных в качестве открытых данных на сайтах-классифайдах, допустимо только с согласия правообладателя[2]. Ярким примером данного утверждения являются дела, одной из сторон которых выступает «Авито» – онлайн-классифайд («доска объявлений»). Так, в деле против сервиса ЦИАН, который обвинялся в копировании части базы объявлений, размещенных первоначально на сервисе «Авито», антимонопольный орган пришел к выводу о нали-

[1] Открытые данные. Роскомнадзор (https://rkn.gov.ru/opendata/).

[2] Подробнее об этом см.: *Рожкова М.А.* Базы данных и сервисы онлайн-классифайдов: пользование базой и использование информации // Журнал Суда по интеллектуальным правам. 2019 (URL: http://ipcmagazine.ru/legal-issues/databases-and-services-online-classification-use-of-the-database-and-use-of-information).

чии в действиях ЦИАН недобросовестной конкуренции, запрещенной ст. 14.8 Федерального закона «О защите конкуренции»[1].

В развитие сказанного нужно признать, что наибольшее коммерческое значение сегодня усматривается за данными, которые в зарубежной литературе обычно называют информационными следами (англ. *footprint data*)[2], понимая под ними всевозможные данные, появляющиеся вследствие действий разных частных лиц (в том числе это может быть информация об определенных личностях, их предпочтениях, местонахождении и т.д.). Такие данные могут поступать из всех перечисленных выше источников, а их обработка на основе *big data analytics* позволяет компаниям реализовывать эффективный таргетинг потребителей, оптимизировать свою деятельность, применять различные управленческие и маркетинговые решения.

При этом «информационные следы» подразделяют на две категории данных: во-первых, *offline footprint data* (данные, накапливаемые вследствие их активности в реальном мире: например, данные о покупках, совершаемых в конкретном магазине, данные об оплате такси кредитной картой) и, во-вторых, *digital (online) footprint data* (так называемый цифровой след, т.е. данные, отображающие действия пользователя в онлайн-среде: в отношении него Рональд Хадкинс (*Ronald E. Hudkins*) подчеркивает, что этим термином обозначается совокупность сведений, которые люди размещают о себе либо оставляют при взаимодействии с различными сайтами в сети Интернет[3]).

В свою очередь «цифровой след» (*digital footprint data*) условно делится на две разновидности: *active digital footprint data* («активный цифровой след»[4], который возникает при размещении пользователями

[1] Антимонопольный орган вынес предупреждение о необходимости прекращения копирования объявлений в сфере недвижимости, размещенных на интернет-сайте *www.avito.ru*, и перенесения скопированных объявлений на интернет-сайт *www.cian.ru*, а также принятия мер по устранению последствий такого нарушения путем удаления ранее скопированного и размещенного на интернет-сайте *www.cian.ru* контента (предупреждение № ИА/9397/18 ФАС России в адрес ООО «Айриэлтор» (сайт ЦИАН) от 13.02.2018).

[2] *Paterson Moira, McDonagh Maeve.* Data protection in an era of big data: The challenges posed by big personal data // Monash University Law Review. Vol. 44. No.1, November 2018. P. 1–31 (https://www.monash.edu/__data/assets/pdf_file/0009/1593630/Paterson-and-McDonagh.pdf).

[3] *Hudkins Ronald E.* Your Digital Footprint Password Protection Requirements (https://www.scribd.com/book/230559848/Your-Digital-Footprint-Password-Protection-Requirements).

[4] В российской публицистике данные, составляющие «активный цифровой след», обычно называют пользовательскими данными, понимая их как составляющую часть

каких-либо данных в Интернете – например, в соцсетях) и *passive digital footprint data* («пассивный цифровой след», который состоит из данных, оставляемых пользователями неосознанно, неспециально (например, информация о посещении интернет-сайта, об использовании бонусной карты в интернет-магазине), и собирается соответствующими программами).

Одно из важнейших последствий создания «*активного цифрового следа*» состоит в том, что пользователь, размещая конкретные данные, касающиеся его личной жизни, в открытом доступе, делает эти данные общедоступными, что исключает для этого лица возможность запрещать использование указанной информации без его согласия. Так, п. 1 ст. 152[2] ГК РФ допускает возможность использования такой информации иными лицами без согласия субъектов этой информации в том числе и в случаях, если информация о частной жизни гражданина ранее стала общедоступной либо была раскрыта самим гражданином или по его воле. При этом вполне ожидаемыми становятся правовые коллизии, поскольку в большинстве случаев «*активный цифровой след*» представляет собой личные сведения (персональные данные), которые подпадают под действие законодательства о персональных данных. В этом большинство экспертов и видят основное препятствие для установления специального правового режима *big data*.

На основе изложенного можно сделать вывод, что особую значимость в контексте *big data* приобретает градация данных на *персональные* и *неперсональные*.

Персональные данные в составе *big data*

Личная информация граждан (англ. *personal data*; в терминологии действующего российского законодательства – персональные данные) в современных реалиях приобрела крайне важное значение: эти сведения ценны с коммерческой точки зрения и одновременно представляют собой существенное благо для каждого конкретного индивида. В связи с этим выделяют два основных подхода при уста-

пользовательского контента (другая составляющая пользовательского контента – это охраняемые объекты интеллектуальной собственности). При этом под пользовательскими данными, по сути, понимают обнародование гражданами в Интернете различной информации о себе (личных сведений): о различных предпочтениях, интересах и склонностях гражданина, которые он считает нужным раскрыть в сети, – любимых фильмах, предпочитаемом бренде одежды, музыкальных пристрастиях и т.д.

новлении правового режима личной информации: 1) ее понимают как товар, объект имущественных прав либо 2) ее трактуют как неотчуждаемое благо, составляющее право человека на неприкосновенность частной жизни[1].

Классическим примером реализации первого подхода являются США, где информация в целом и личная информация рассматриваются как товар (англ. *commodity*). Этот подход не означает полное отрицание прав индивидов на охрану частной жизни, но общий вектор направленности законодательства состоит в обеспечении коммерциализации личной информации и ее свободного оборота. Такой подход выгоден для компаний, работающих с новыми технологиями и, в особенности, с big data. Но в то же время он оставляет открытыми вопросы относительно защищенности граждан от вмешательства в их частную жизнь.

Показательным здесь является скандал вокруг *Cambridge Analytica* – аналитической компании, которая с помощью *big data analytics* собирала и обрабатывала данные пользователей *Facebook* без какого-либо согласия пользователей, а затем использовала полученные результаты, в том числе с целью влияния на политические предпочтения пользователей[2]. Этот случай стал катализатором ужесточения в США законодательства в сфере защиты данных и, в частности, повлек принятие в штате Калифорния Акта о конфиденциальности данных (англ. *The California Consumer Privacy Act*), который вступил в силу 01.01.2020. Этот Акт нацелен на защиту прав граждан при обработке их личной информации и во многом копирует передовые механизмы европейского законодательства по защите данных (о них будет говориться далее). Вместе с тем калифорнийский закон не отступает от постулата «данные – товар»: защита прав индивида предполагается в рамках коммерческих отношений, в том числе по продаже личной информации. Здесь же надо отметить, что в настоящее время и другие штаты в США начали разрабатывать собственные акты о защите личной информации, иногда звучат предложения и о разработке в США федерального закона в данной сфере[3].

---

[1] См., например: *Spilman Maia T.* EACC Insights: Personal Data: a Commodity or a Right? (https://www.eaccny.com/news/eacc-insights-personal-data-a-commodity-or-a-right/#_ftn1).

[2] Facebook-Cambridge Analytica scandal // https://www.bbc.com/news/topics/c81zyn0888lt/facebook-cambridge-analytica-scandal

[3] США хотят скопировать Закон ЕС о защите персональных данных, но лишь частично (URL:https://www.vestifinance.ru/articles/119700 (дата обращения: 01.07.2019)).

Другой подход к вопросу защиты личной информации сложился в европейском правопорядке – он основан на основополагающих актах, закрепляющих фундаментальные права человека, включая право человека на неприкосновенность частной жизни. Среди таких актов Всеобщая декларация прав человека 1948 г., Международный пакт о гражданских и политических правах человека 1976 г., Европейская конвенция о защите прав человека и основных свобод 1950 г. и Конвенция Совета Европы о защите физических лиц при автоматизированной обработке персональных данных (*ETS № 108*) 1981 г. (далее – Конвенция 108).

Конвенция 108 заслуживает особого внимания в рамках настоящей статьи, поскольку представляет собой международный договор, нацеленный на решение проблем, порождаемых использованием новых технологий и обеспечение фундаментальных прав человека при автоматизированной обработке информации.

В преамбуле Конвенции 108 отмечается увеличение трансграничного потока личной информации (*personal data*), которая сегодня подвергается автоматизированной обработке, включающей хранение, аналитику, изменение, уничтожение, поиск, распространение личных сведений. Указано, что в этих условиях необходимо усиление защиты прав и свобод граждан, в частности права на неприкосновенность частной жизни. С учетом этого цель Конвенции 108 состоит в обеспечении для каждого частного лица уважения его прав и основных свобод, в частности его права на неприкосновенность частной жизни в том, что касается автоматизированной обработки его личной информации (ст. 1). Для краткости эта цель обозначена как «защита данных». Иными словами, использованное в Конвенции 108 выражение «**защита данных**» надо понимать в контексте защиты не самих по себе личных данных, а прав и основных свобод частных лиц, которые могут быть нарушены при автоматизированной обработке (использовании) этой информации.

В рамках процесса подготовки ратификации Россией Конвенции 108[1] был принят Закон о персональных данных. Цель данного Закона заключалась в том числе в приближении по содержанию охраняемых прав и свобод человека к европейским стандартам. Закон о персональных данных отражает большинство положений Конвен-

[1] По сообщению МИД России 15.05.2013 Россия завершила процедуру ратификации Конвенции СЕ о защите физлиц при автоматизированной обработке персональных данных (URL: https://www.garant.ru/products/ipo/prime/doc/70281462/).

ции 108 и, как указано в его ст. 2, преследует цель обеспечения защиты прав и свобод человека и гражданина при обработке его персональных данных, в том числе защиты прав на неприкосновенность частной жизни, личную и семейную тайну.

В то же время само выражение «персональные данные», которым оперирует Закон о персональных данных, сложно назвать удачным. В Конвенции 108 используется уже упоминавшийся термин *personal data*, который на русский язык точнее было бы перевести как «личные данные», «личные сведения» или «личная информация». Это связано с тем, что любой из перечисленных вариантов указывает на связь информации с личностью, на частный характер этой информации. Введенный в отечественное законодательство термин «персональные данные» такой посыл не несет и подспудно воспринимается как информация, которая собирается, используется и контролируется государственными органами. Вместе с тем, учитывая, что термин «персональные данные» уже закреплен отечественным законодательством и прочно вошел в правовой оборот, авторы настоящей статьи не предлагают заменить его другим, но в настоящей статье используют термины «персональные данные», «личная информация» и «личные сведения» как синонимы.

Дефиниция «персональные данные», содержащаяся в российском законе, по сути повторяет определение, изложенное в Конвенции 108: «...любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу». Данное определение является весьма широким, скорее рамочным. В доктрине и правоприменительной практике можно встретить разнообразные попытки на основе данного понятия выделить конкретные критерии для отнесения информации к персональным данным.

Так, достаточно детальный анализ понятия *personal data* был проведен Рабочей группой Статьи 29 (англ. *The Article 29 Working Party*)[1]. Это аналитическое исследование проводилось в преломлении европейского законодательства и относительно давно, однако отраженный в нем подход к содержанию понятия «персональные данные» в общем-то не поменялся (как известно, подход, формирующийся в российском законодательстве, тяготеет к европейскому). В своем исследовании Рабочая группа выделила четыре фундаментальных блока, на кото-

[1] Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data (2007) (https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm).

рых строится понимание персональных данных: 1) информация (разнообразные категории, различная форма фиксации, необязательно достоверность и т.д.); 2) связь между информацией и субъектом (выделяются три элемента связи: содержание, задачи и интерпретация результатов); 3) определенный или определяемый субъект (существует разумная вероятность, что на основе информации можно идентифицировать определенного субъекта); 4) субъект персональных данных (физическое лицо).

На практике наиболее неоднозначный и сложно определяемый критерий – это возможность идентификации. В связи с этим Л. Детерманн пишет: «...не обязательно, чтобы такие данные сами по себе позволяли идентифицировать субъекта данных. Достаточно того, чтобы данные в разумной степени относились бы к физическому лицу, которое можно идентифицировать»[1]. Данная идея прослеживается и в упомянутом исследовании Рабочей группы Статьи 29, в котором отмечается, что вовсе не обязательно, чтобы личная информация сама по себе являлась достаточной для того, чтобы точно определить конкретного индивида. Требуется, чтобы при сопоставлении данной информации с другой, разумно доступной, была *возможность выделить* субъекта персональных данных из группы иных лиц. Например, в Интернете имеется возможность получить большое количество информации о социоэкономических, психологических и иных чертах субъекта и составить образ личности без знания прямых идентификаторов – таких, как имя, адрес проживания и т.д. Основываясь на данных рассуждениях, можно получить однозначный ответ, почему, например, *IP*-адрес относится к персональным данным.

В ряде случаев законодательство прямо указывает, что та или иная информация относится к персональным данным.

Так, новый европейский закон в сфере защиты данных – *GDPR* предусматривает, что к персональным данным в том числе относятся: имя, данные о местоположении, онлайн-идентификаторы, а также факторы, характерные для физической, физиологической, генетической, умственной, экономической, культурной или социальной идентичности конкретного физического лица[2]. Но этот перечень, что очевидно, не является закрытым – исчерпывающим образом пере-

[1] *Детерманн Л.* Путеводитель в правовом регулировании персональных данных Лотара Детерманна Международный корпоративный комплаенс. М.: Инфотропик, 2018. С. 15.

[2] Art. 4 GDPR Definitions // https://gdpr-info.eu/art-4-gdpr/

числить все персональные данные не представляется возможным: под режим персональных данных (как по *GDPR*, так и по российскому законодательству) могут подпадать разнообразные виды личной информации, которые позволяют идентифицировать человека.

Надо заметить, что основные принципы обработки персональных данных, закрепленные в *GDPR*, значительно повышают стандарты охраны прав субъектов персональных данных, нежели установленные Конвенцией 108. *GDPR* вводит ряд существенных требований к обработке персональных данных, направленных на повышение уровня защиты субъектов персональных данных, дополнительно закрепляет право субъектов персональных данных на перенос данных от одного оператора к другому, уточняет содержание права на доступ к своим персональным данным и их удаление и т.д. Причем, что немаловажно с практической точки зрения, за нарушение законодательства о персональных данных *GDPR* вводит действительно существенные штрафы – за нарушения до 4% глобального оборота или 20 млн евро.

Более того, штрафы за нарушение *GDPR* активно применяются. Статистика показывает, что с момента вступления в силу *GDPR* регуляторы взыскали штрафов на общую сумму 114 млн евро, зарегистрировав 160 тыс. нарушений[1]. Например, в июне 2019 г. мебельная компания *IDdesign*, допустившая нарушение порядка хранения персональных данных (после достижения цели обработки данные клиентов не удалялись), была оштрафована на 200 тыс. евро, а французская компания в сфере недвижимости *Sergic* за несоблюдение порядка хранения данных в совокупности с нарушениями требований безопасности хранения данных была оштрафована на 400 тыс. евро[2]. Самый крупный штраф в размере 50 млн евро был возложен на *Google*.

Некоторые тенденции в сторону ужесточения законодательства о защите данных ожидаются и в России, это связано с подписанием Россией Протокола к Конвенции 108, который существенно изменяет содержание этого международного договора, вследствие чего Конвенцию в редакции данного Протокола нередко именуют как «Конвенция 108+».

«Конвенция 108+» закрепляет ряд положений, основанных на *GDPR*, и предъявляет более высокие требования к обработке и защите данных, нежели к содержавшимся в прежней редакции

[1] https://habr.com/ru/news/t/485320/

[2] Data Protection News & Resources June 2019 (https://www.evalian.co.uk/data-protection-news-resources-june-2019/).

Конвенции 108. В частности, вводятся новые категории особых персональных данных (генетические данные, сведения о членстве в профсоюзах и этническом происхождении); появляется обязанность оператора персональных данных незамедлительно направлять в орган по защите данных уведомление об утечке данных; вводятся дополнительные гарантии для субъектов персональных данных и новые права (например, право получать информацию о процессах, лежащих в основе обработки, право заявлять возражения)[1]. Ожидается, что «Конвенция 108+» позволит гармонизировать законодательство в сфере персональных данных различных юрисдикций, взяв за основу европейский опыт. Однако речь идет скорее об обеспечении так называемого адекватного уровня защиты данных, а не о полном совпадении законодательства.

При этом нельзя не замечать, что отличия национальноого законодательства о защите данных от европейского могут корениться уже в самой терминологии. В частности, Л. Детерманн обращает внимание на то, что в отличие от европейского законодательства, в котором находит применение только понятие «персональные данные», в законодательстве других стран употребляются и иные термины: «Так, в нескольких разделах Гражданского кодекса Калифорнии используется понятие «персональная идентифицирующая информация»; данный термин определен узко и в каждом случае по-разному, в зависимости от контекста и цели того или иного раздела»[2]. При этом автор подчеркивает и различие в понимании *специальных категорий* персональных данных: «Например, в США особой защите подлежат номера Федеральной программы социального страхования, присваиваемые каждому гражданину, и данные кредитных карт, поскольку потребители часто становятся жертвами краж идентифицирующих данных пользователей из-за относительно слабых процедур аутентификации, используемых банками и продавцами. Кража таких данных является предметом меньшего беспокойства в Европе, однако европейские компании по общему правилу должны соблюдать особые ограничения в отношении персональных данных, касающихся: политических взглядов; членства в профсоюзах; состояния здоровья (например, больничные

[1] Информационное письмо АЛРУД «Предстоящие существенные изменения российского законодательства в области защиты персональных данных» от 07. 11. 2018 (http://www.alrud.ru/upload/iblock/080/Информационное%20письмо_Предстоящие%20существенные%20изменения%20российского%20законодательства%20в%20области%20защиты%20персональных%20данных.pdf).

[2] *Детерманн Л.* Указ. соч. С. 17.

дни сотрудника, рецепты выписанных лекарств, данные медицинских тестов, даже если они хранятся с использованием идентификаторов без указания имен пациентов); расового или этнического происхождения (например, место рождения, фотографии, указывающие на цвет кожи); религиозных или философских убеждений (например, налоговый статус сотрудника для целей начисления церковного налога в Германии); информации, относящейся к сексуальной ориентации (например, семейное положение в странах, которые не признают однополые браки); некоторых типов судимости»[1].

Исходя из сказанного следует обратить внимание на то обстоятельство, что вопрос использования личной информации может регулироваться различными нормативно-правовыми актами, предписывающими определенный порядок такого использования. Однако с учетом того, что такая информация является личной, на нее распространяется и законодательство о персональных данных.

Анализ российского законодательства позволяет выделять, в частности, следующие разновидности данных, которые следует относить к персональным:

– *уникальные идентификаторы человека*, в частности имя, отчество, фамилия гражданина, а также его псевдоним, в том числе творческий. Возможность использования этих уникальных идентификаторов с согласия гражданина другими лицами в их творческой, предпринимательской и иной экономической деятельности прямо предусмотрена законом (п. 4 ст. 19 ГК РФ);

– *изображение гражданина* (в виде фотографий, видеозаписей и проч.), представляющее собой разновидность информации о самом лице или его частной жизни. Каждое физическое лицо является обладателем абсолютных прав на собственное изображение и вправе как само использовать свое изображение (например, размещая свои фотографии в *Instagram*), так и распоряжаться правом на свое изображение (в частности, предоставляя заинтересованным лицам право публикации своих фотографий в различных изданиях). И, например, согласно п. 1 ст. 152.1 ГК РФ допускается использование изображения гражданина после его смерти супругом и детьми, а при их отсутствии – родителями;

– *уникальные идентификационные номера и иные идентифицирующие сведения*, к которым традиционно относят, в частности, ИНН, СНИЛС, присваиваемые в рамках различных государственных систем,

[1] *Детерманн Л.* Указ. соч. С. 18.

а также паспортные данные, данные водительских удостоверений и проч. Учитывая, что подобные номера необходимы для достижения общественно значимых (публичных) целей, исключена возможность признания на них имущественных прав граждан, но вместе с тем граждане являются субъектами этой информации, поэтому за ними следует признавать неимущественные права на эту информацию;

– *особо значимые конфиденциальные личные сведения* (называемые иногда особо чувствительными данными), к которым сегодня принято относить сведения в финансовой сфере, о здоровье, детях, о социальном страховании, а также геолокационные данные[1].

Является очевидным, что состав персональных данных чрезвычайно многообразен и весьма велика вероятность того, что в дальнейшем это разнообразие будет только увеличиваться. Причем как «уже признанные», так и новые персональные данные будут представлять колоссальный коммерческий интерес, в том числе для целей их использования в *big data analytics*.

Основная проблема «попадания» персональных данных в состав *big data* и их использования при осуществлении *big data analytics* связана с тем, что законодательство о персональных данных применительно к обработке данных предусматривает, во-первых, необходимость *получения согласия* субъектов данных на обработку их персональных данных, во-вторых, ограничение обработки *заранее определенными* целями, совместимыми с целями сбора персональных данных, в-третьих, ограничение содержания и объема обрабатываемых данных *заявленным целям обработки*. Вместе с тем, как указывалось выше, *big data analytics* предполагает многократное использование доступных наборов данных, причем нередко для целей, отличающихся от тех, для которых эти данные первоначально предназначались. Следует подчеркнуть, что здесь речь не идет о неправомерных целях: данные могли изначально собираться, например, для целей изучения книжных пристрастий широкого круга читателей, а впоследствии использоваться для решения маркетинговых задач по продвижению *e-book* (при этом читатели, давшие согласие на использование своих данных для первого случая, на вторичное использование своих данных согласия не давали). Следовательно, при проведении *big data analytics* не соблюдается сразу несколько основных принципов обработки персональных данных.

[1] См. об этом § 3.4.2 статьи М.А. Рожковой «Имущественные права на новые нематериальные объекты в системе абсолютных прав» в настоящем ежегоднике.

В связи с этим возникает немало дискуссий по поводу легальности использования *big data* и потенциальной несовместимости как *GDPR*, так и национального законодательства о персональных данных с бизнес-решениями, построенными на *big data analytics*[1].

Однако, на наш взгляд, требования, предъявляемые сегодня как *GDPR*, так и национальным законодательством, не являются непреодолимым барьером для коммерциализации *big data* и использования в *big data analytics*.

В частности, *GDPR* достаточно гибко подходит к законным основаниям обработки данных, причем распространенным основанием является законный интерес контроллера («оператора» в терминологии российского законодательства). Наличие законного интереса у контроллера определяется с помощью трех тестов на наличие: 1) законной цели, 2) необходимости, 3) соблюдения баланса интересов. При соответствии этим критериям и с опорой на законный интерес компании могут обрабатывать персональные данные в составе *big data*, например, для повышения качества, стабильности сервиса без получения отдельного согласия пользователя. Но стоит учитывать, что наличие законного основания – это только одно из требований законодательства и его соблюдение не освобождает компании от иных условий правомерной обработки данных, включая обязанность сообщения субъекту персональных данных до сбора его данных о целях и законном интересе обработки, предоставлении субъекту персональных данных права на возражение против обработки, права на удаление и ограничение.

Например, при осуществлении маркетинговой деятельности компания может собирать и обрабатывать персональные данные, которые размещены в открытых источниках в сети Интернет[2], с целью улучшения отдельных функций сервиса, персонализации их для клиентов.

[1] См., например: *Zarsky Tal*. Incompatible: The GDPR in the Age of Big Data // Seton Hall Law Review. August 8, 2017. Vol. 47. No. 4 (2) (https://ssrn.com/abstract=3022646).

[2] Стоит отметить, что *GDPR* также конкретизирует режим персональных данных, получаемых опосредованно через третьих лиц. Такие данные остаются персональными, при их обработке необходимо соблюдение требований законодательства о персональных данных. Однако в *GDPR* разработан механизм их использования: прописаны конкретные обязанности контроллера по нотификации субъекта персональных данных, предусмотрен перечень исключений из общих правил, в том числе связанных с разумной невозможностью нотификации конкретного субъекта персональных данных, а также с наличием законного интереса контроллера на обработку таких данных (GDPR. Article 14. Information to be provided where personal data have not been obtained from the data subject).

Компания может обосновать, что достижение поставленной цели невозможно или существенно затруднено без обработки персональных данных, обработка носит очевидный и открытый для субъекта характер, не наносит или разумно не должна приносить вред субъекту персональных данных. То есть в любом случае требуется информирование субъекта, соблюдение принципов и основных условий обработки персональных данных.

Однако в ряде случае осуществление таких действий не представляется возможным. В таких ситуациях наиболее перспективным механизмом будет являться анонимизация персональных данных, позволяющая превратить персональные данные в неперсональные.

Анонимизированные данные: на пути от персональных к неперсональным

Анонимизация персональных данных представляет собой метод, который позволяет удалить идентифицирующую конкретного индивида информацию из массива данных[1]. Согласно позиции Рабочей группы Статьи 29 для соблюдения стандартов анонимизациии персональных данных необходимо, чтобы данные были отделены от любых идентификаторов и не позволяли более определять конкретного субъекта персональных данных[2]. Иными словами, анонимизация предполагает, что данные обработаны таким образом, который исключает вероятность последующей идентификации физического лица с помощью любых средств, которые разумно могут применяться третьими лицами.

Примечательно, что наряду с анонимизацией существует и **псевдонимизация**, которая в GDPR трактуется как самостоятельный механизм, разработанный для целей защиты персональных данных[3]. Псевдонимизация по *GDPR* является аналогом российского механизма

[1] И. де Монжуа, научный сотрудник Массачусетского технологического института, объясняет анонимизацию данных как двухэтапный процесс, включающий псевдонимизацию (англ. *pseudonymization*) и деидентификацию (англ. *de-identification*). См.: *Montjoye Yves-Alexandre de*. Data anonymization techniques less reliable in era of big data // https://searchcompliance.techtarget.com/feature/High-dimensional-info-complicates-data-anonymization-techniques.

[2] Working Party Article 29-Opinion 05-2014 on Anonymisation Techniques (https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm).

[3] См., например: Top 10 operational impacts of the GDPR: Part 8 – Pseudonymization // https://iapp.org/news/a/top-10-operational-impacts-of-the-gdpr-part-8-pseudonymization/

обезличивания персональных данных, т.е. действий, результатом которых становится невозможность без использования дополнительной информации определить принадлежность персональных данных конкретному лицу. Классический пример обезличивания – использование методов шифрования, когда определенные идентификаторы (имя и фамилия, дата рождения и т.д.) заменяются, например, буквенным шифром. Используемый шифр не позволяет установить конкретного человека, но только до момента, пока не получены ключи шифрования, или не выявлен принцип, с помощью которого данные были зашифрованы. Обезличенные (псевдонимизированные) данные являются персональными, и к их обработке применяются общие требования закона о персональных данных.

В отличие от обезличивания (псевдонимизации) анонимизация персональных данных всегда строится на **деиндентификации**, которая представляет собой необратимый процесс, окончательно исключающий возможность установить конкретного человека на основе некоторой совокупности данных. Информация, подвергшаяся анонимизации, перестает относиться к персональным данным, т.е. анонимизация превращает персональные данные в неперсональные, в результате чего предоставляется большая свобода их использования, в том числе и в рамках *big data analytics*. В зарубежной практике выработаны определенные стандарты и методики анонимизации персональных данных, ряд которых одобрен регуляторами в сфере персональных данных[1].

В то же время все чаще в научной литературе представляется позиция, согласно которой анонимизация персональных данных – это не более чем фикция: современный уровень технологического развития позволяет «деанонимизировать» практически любые данные, тем самым повторно «реидентифицировать» конкретного человека[2]. Иными словами, анонимизация не может окончательно и бесповоротно исключить возможности сопоставления полученных данных. Зачастую подвергается критике сама возможность исчерпывающего определения перечня идентификаторов, которые необходимо

[1] Anonymisation: managing data protection risk code of practice. ICO // (https://ico.org.uk/media/1061/anonymisation-code.pdf

[2] *Ohm Paul.* Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). UCLA Law Review, Vol. 57. P. 1707, 2010*; Narayanan A. Shmatikov V.* 2007. How to break the anonymity of the Netflix Prize dataset (URL: https://arxiv.org/abs/cs/0610105; Pete Warden. Why you can't really anonymize your data (https://www.oreilly.com/ideas/anonymize-data-limits)).

изъять, чтобы по полученным данным стало однозначно невозможно реидентифицировать первоначального субъекта персональных данных.

Кроме того, ряд авторов критически подходят к анонимизации исходя из потери ценности персональных данных после такой обработки[1]. С такими утверждениями сложно спорить, поскольку, например, для прямого таргетинга действительно ценна именно связь данных с конкретной личностью. Но в то же время нельзя не учитывать, что анонимизированные данные все-таки не являются абстрактной статистической информацией – такие данные могут представлять значительную информационную и коммерческую ценность и использоваться для решения конкретных бизнес-задач. Например, для принятия компанией решения о расширении производства в определенном сегменте решающую роль будут играть данные об общем уровне спроса на те или иные товары и услуги, информация о предпочтениях и моделях поведения потребителей – в идентификации субъектов персональных данных здесь нет никакой необходимости. Другим наглядным примером ценности анонимизированных данных является их применение транспортными компаниями для расчета оптимальных маршрутов, определения временных интервалов в расписании и т.д.

Таким образом, несмотря на некоторые несовершенства процедуры анонимизации персональных данных, сложности ее осуществления с технической точки зрения и потерю определенных возможностей в использовании получаемой информации, сами анонимизированные данные представляют собой ценный нематериальный актив. С юридической точки зрения анонимизацию данных следует рассматривать как инструмент, который позволяет превратить персональные данные в неперсональные, тем самым снимая значительное количество ограничений, связанных с законодательными требованиями к обработке персональных данных, и разрешая проблемы, связанные с неотчуждаемостью определенных категорий личной информации, входящей в понятие «персональные данные», обремененностью персональных данных личными неимущественными, а в определенных случаях и имущественными правами конкретных лиц.

[1] *Mattioli Michael*. Disclosing Big Data // Minnesota Law Review. Vol. 99. No. 2. February 20, 2014. P. 566.

Неперсональные данные

Говоря о *big data*, конечно, невозможно ограничиваться лишь вопросами, связанными с персональными данными – не меньшего внимания заслуживают так называемые неперсональные данные.

Термин «**неперсональные данные**» активно используется в европейском законодательстве. В Европейском союзе даже принят специальный Регламент 2018/1807 Европейского парламента и Совета ЕС от 14.11.2018 по структуре свободного обращения неперсональных данных в Европейском союзе (далее – Регламент неперсональных данных). Данный Регламент регулирует вопросы перемещения неличных данных через границу и обеспечения свободы в предоставлении услуг по обработке данных в пределах Европейского Союза.

В Регламенте отсутствует дефиниция понятия «неперсональные данные» – к их числу предлагается относить все данные, не являющиеся персональными. С учетом этого к неперсональным данным можно причислять не только анонимизированные данные, но и все данные, поступающие из *технических источников*.

Это, в частности, данные, создаваемые в рамках упомянутого интернета вещей, включающего в свой состав и промышленный интернет, который «поставляет» информацию со всевозможных датчиков, котроллеров, приборов учета потребления, устройств аудио- и видеорегистрации, измерительных комплексов и проч. Сюда же будут включены данные из таких источников, как искусственный интеллект (англ. *artificial intelligence*) и машинное обучение (англ. *machine learning*). К неперсональным данным относят и такие востребованные практикой категории данных, как метеорологические, экологические и географические сведения, данные финансовых и страховых рынков, государственная статистика и информация из государственных реестров (не касающаяся субъектов персональных данных), экономические показатели, результаты аналитических и научных исследований.

Примечательно, что некоторые наборы или базы неперсональных данных могут размещаться в форме открытых данных. И поскольку такие данные не привязаны к субъекту, то их применение для целей *big data analytics*, казалось бы, не предполагает получение какого-либо предварительного согласия на их использование. Между тем это не совсем так.

Неперсональные данные возникают не сами по себе – они являются результатом деятельности конкретных лиц по их сбору, накоп-

лению, структуризации и иной обработке. Вследствие сказанного лица, осуществившие подобную деятельность, могут рассматриваться как правообладатели такой информации, что предполагает получение у них согласия на использование неперсональных данных (данный вопрос требует самостоятельного изучения и не раскрывается в данной работе).

Но можно отметить, что в ряде случаев такой подход мы уже наблюдаем на практике: массивы данных Федеральной службы по гидрометеорологии и мониторингу окружающей среды[1], базы данных с аналитическими материалами по инвестиционному рынку[2], разнообразные базы судебных решений и иных правоприменительных актов.

Заключение

Завершая настоящую работу, нельзя не вспомнить, что *big data* называют новой нефтью. Но смысл фразы, прозвучавшей в 2006 г. и приписываемой британскому математику Клайву Хамби[3]: «Данные – это новая нефть!» (англ. *Data is the new oil!*), – не в признании равноценности данных и нефти, а в констатации того, что, как и сырая нефть, необработанные (сырые) данные не представляют собой особой ценности: для того чтобы данные создавали действительную прибыль, они должны быть использованы – подвергнуты анализу, визуализированы, интегрированы и т.д. для целей принятия того или иного решения. Следовательно, на первый план выходит понимание *big data* не как огромных объемов информации, а как новых технологий, связанных с обработкой этой информации, и прежде всего big data analytics, позволяющего принимать коммерчески значимые решения, развивать производство, оценивать риски.

Вместе с тем в качестве своего рода «помехи», препятствующей активной экспансии *big data analytics*, сегодня рассматривают персональные данные, которые, являясь необходимым механизмом защиты прав человека в условиях современного общества, плохо сочетаются

[1] Специализированные массивы (http://meteo.ru/data).

[2] Подписка на аналитические материалы. *The Wall Street professional* (https://thewallstreet.pro/subscription).

[3] *Palmer Michael*. Data is a new oil // https://ana.blogs.com/maestros/2006/11/data_is_the_new.html

с идеей использования *big data* и практического внедрения *big data analytics*. Эта проблема носит глобальный характер и сейчас идет активная разработка различных вариантов компромисса между потребностями инновационного развития и защиты права человека на частную жизнь.

Пристатейный библиографический список

1. *Детерманн Л.* Путеводитель в правовом регулировании персональных данных Лотара Детерманна Международный корпоративный комплаенс. М.: Инфотропик, 2018.

2. *Рожкова М.А.* Характеристики больших данных, значимые для целей гражданского права // Хозяйство и право. 2019. № 6.

3. Hudkins Ronald E. Your Digital Footprint Password Protection Requirements, June 18, 2014. ISBN 9781500192631 (https://www.scribd.com/book/230559848/Your-Digital-Footprint-Password-Protection-Requirements).

4. *Mattioli Michael*. Disclosing Big Data // Minnesota Law Review, Vol. 99. No. 2. February 20, 2014, 2014.

5. *Mayer-Schönberger V. and Cukier K.* Big Data: A Revolution that Will Transform how We Live, Work, and Think. Houghton Mifflin Harcourt, 2013.

6. *Montjoye Yves-Alexandre de*. Data anonymization techniques less reliable in era of big data // https://searchcompliance.techtarget.com/feature/High-dimensional-info-complicates-data-anonymization-techniques

7. *Narayanan A., Shmatikov V.* How to break the anonymity of the Netflix Prize dataset. 2007 // https://arxiv.org/abs/cs/0610105

8. *Ohm Paul* . Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization // UCLA Law Review. Vol. 57. August 13, 2009.

9. *Paterson Moira Paterson, McDonagh Maeve McDonagh.* Data protection in an era of big data: The challenges posed by big personal data. // Monash University Law Review., Vol. No. 1. November 2018.

10. *Spilman Maia T.* EACC Insights: Personal Data: a Commodity or a Right? (https://www.eaccny.com/news/eacc-insights-personal-data-a-commodity-or-a-right/#_ftn1).

11. Warden Pete . Why you can't really anonymize your data // (ttps://www.oreilly.com/ideas/anonymize-data-limits

12. Working Party Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data (2007) // https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm

13. Working Party Article 29-Opinion 05-2014 on Anonymisation Techniques // https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm

14. *Zarsk, Tal*. Incompatible: The GDPR in the Age of Big Data // Seton Hall Law Review. Vol. 47, No. 4 (2). August 8, 2017.