DOI: <http://dx.doi.org/10.15688/jvolsu7.2016.1.7>

УДК 177.9:002

ББК 87.713.12

СПРАВЕДЛИВОСТЬ В ИНФОРМАЦИОННОЙ ОБЛАСТИ: ЭТИЧЕСКИЙ АСПЕКТ BIG DATA ¹

Евгений Владимирович Карчагин

Кандидат философских наук,
доцент кафедры философии, социологии и психологии,
Волгоградский государственный архитектурно-строительный университет
evgenkar@yandex.ru
ул. Академическая, 1, 400074 г. Волгоград, Российская Федерация

Аннотация. В статье рассматривается проблема справедливости применительно к современной ситуации, обусловленной появлением нового источника и метода обработки информации – Big Data («Большие Данные»). Среди основных источников необходимости переосмыслить справедливость в информационной области автором выделяются: персональные данные и их приватный характер, вступающий в конфликт с современными требованиями безопасности и практиками упреждающего слежения; общественные движения за открытость данных; неравный доступ к Большим Данным, порождающий новый тип социальной несправедливости – цифровое неравенство, выражающееся в расслоении из-за неравенства в доступе к информационным технологиям; проблема дискриминации и использование Больших Данных.

Ключевые слова: Big Data, справедливость, этика, информационная область справедливости, приватность и публичность данных, цифровое неравенство.

Появление Big Data, Больших Данных (далее – БД) вызвало целую волну дискуссий и публикаций, осмысляющих новые возможности в научной сфере, в том числе в социальных науках. Одной из главных линий изучения БД стала линия эпистемологического и методологического анализа. Другая линия – не менее важная – состоит в осмыслении этических перспектив использования БД. В частности, феномен БД может выступить как иллюстрация или даже источник для выявления и демонстрации проблемы справедливости, понимаемой в широком смысле как идеи оправданности того или иного порядка в распределении различных социальных благ и тягот, в информационной области. Конечно, БД не всегда и не во всех своих аспектах могут соотноситься с этическими вопросами, но только если эти данные имеют отношение к соци-

альному миру: когда это напрямую данные о людях, либо иные данные, но используемые таким способом, который может пониматься как несправедливый и неэтичный. Целью статьи является раскрытие ключевых аспектов проявления проблемы справедливости в информационной области, обусловленной появлением нового источника и способа обработки информации – БД.

Приватность

В первую очередь социальная и этическая проблематичность БД выражается в вопросе о *приватности данных*. Можно различать два направления в современной дискуссии о частном и приватном характере данных: 1) приватность данных пользователей, частных лиц, которые даже зачастую не осведом-

лены о том, что некоторые стороны их жизни конвертированы в данные; 2) приватность собственности данных владельца, в том числе корпораций, университетов и т. д. Если второй вид приватности связан с правовым полем, то второй во многом репрезентирует сложную моральную проблему этической оправданности и справедливости вмешательства в частную жизнь человека.

С увеличением и детализацией массивов данных становится все легче идентифицировать индивида при их помощи. Один из ярких примеров этого имел место, когда два исследователя из Техасского университета смогли идентифицировать определенных индивидов в массивах данных стримингового видео-сервиса Netflix с помощью сравнения анонимных данных из Netflix с обзорами и рецензиями, помещенными на другом сайте, что в итоге привело к судебному иску. Другой пример новых возможностей – исследователи теперь способны идентифицировать человека с опорой на информацию, содержащуюся внутри массива генетических данных, используя не только генетические данные, но и генеалогические веб-сайты, а также публично доступные данные [10, р. 132]. Это, в частности, ставит вопрос о том, как ученые могут работать с большими множествами данных, сохраняя при этом защиту частной жизни.

Приватность, частная жизнь проблематичны и многозначны, при том что они крайне важны для общественного роста и социальных взаимодействий. Невозможно быть социальным субъектом, не обладая приватностью и не имея возможности уединяться. Однако интерпретация приватности не является однозначной, и способна иметь «перегибы». Такие перегибы особенно становятся заметными на примере информационной политики крупных социальных сетей. Так, по поводу негласной позиции в этом вопросе Facebook К. Вайдхьянатан и С. Булок критически восклицают: «Марк Цукерберг считает, что, будучи приватными, будучи не открытыми, не показываемыми всё для всех, мы будем в чем-то неаутентичными и, таким образом, антисоциальными. Он думает, что он определяет, что значит быть социальным» [12, р. 63].

Дана Бойд и Кейт Кроуфорд в своей часто цитируемой статье «Критические вопро-

сы к Большим Данным» указали на принцип нетождественности, доступности и этичности: «только потому что нечто доступно не делает это этичным его использование». Какой статус имеют публичные данные в социальных сетях? Можно ли их использовать без разрешения? «Что если чей-то «публичный» пост выдернут из контекста и проанализируют таким образом, как автор не мог представить? Что значит для кого-то быть изучаемым и анализируемым без всякого ведома? Кто отвечает за уверенность в том, что отдельные лица и сообщества не пострадают от процесса исследования? Как выглядит информированное согласие?» [3, р. 672]. Все эти вопросы сводятся к проблеме контроля данных, когда пользователи социальных сетей не имеют доступа к инструментам контроля за данными и зачастую даже не осведомлены о том, что эти данные используются кем-то: «... вполне возможно, что некоторые пользователи не дали бы разрешение на использование своих данных где-то еще. Многие не осознают множественности агентов и алгоритмов осуществляющих текущий сбор и хранение данных для дальнейшего использования. <...> Данные могут быть публичными (или полупубличными), но это не просто равняется полному разрешению для всех видов их использования. Исследователи Больших Данных редко признают, что существует значительная разница между тем, чтобы быть на публике (то есть сидеть в парке) и быть публичным (то есть активно привлекать внимание)» [3, р. 673].

Слежение

Важная проблема, напрямую связанная с приватностью, – практики *слежения*, то есть «любое систематическое, рутинное и фокусированное внимание к личным подробностям для заданной цели (такой как управление, влияние или наделение правом)» [8, р. 2], осуществляющиеся на базе БД и веб-отслеживания. Веб-отслеживание это – незаметное и несанкционированное извлечение, хранение, анализ, продажа и покупка личных онлайн данных, присвоенное одним или несколькими удаленными корпоративными онлайн-субъектами.

Skype, к примеру, собирает в числе прочего такие типы данных о своих пользователях: личные данные (имя, пользовательское ID, возраст, пол, язык, список контактов); данные о личной собственности (тип компьютера, возможности компьютера, операционная система, тип браузера, модель мобильного телефона, тип и возможности роутера и т. п.); финансовые данные (номер кредитки, банковская информация); данные о поведении (куки, пользовательская статистика, купленные и полученные товары и услуги, статистика просмотра страниц, расстояния звонков, данные о трафике, содержание сообщений, голосовые и видео сообщения, доступ к аккаунтам в Facebook, Microsoft, электронной почте); данные о провайдере (пропускная способность, адрес электронной почты, IP-адрес, отчеты об ошибках и др.) [11, р. 3]. Весьма впечатляющий список.

С. Пикок отмечает серьезность вопроса о приемлемости веб-отслеживания в контексте появления БД: «Нынешний неравный онлайн-обмен осуществляется не между двумя полностью информированными агентами и он не улучшает публичную онлайн-сферу. Он может быть схвачен, правда, другим термином, гораздо менее лестным, чем компромисс: он имеет характеристики нечестного договора. Нечестные договоры являются эксплуататорскими, несправедливыми, безальтернативными и перекладывают бремя экономической сделки целиком на одну сторону, и в этом случае это интернет-пользователь. Для большинства пользователей внутренне несправедливое и неравное бремя нигде не прописано, и даже если бы было иначе, люди вряд ли могли бы ознакомиться с условиями. В более юмористической версии этого серьезного вопроса, Huffington Post сообщила, что в 2010 г. около 7 500 онлайн-пользователи «случайно продали свои души» игровому интернет-магазину» [11, р. 7].

В последнее время практики слежения перешли от целевого контроля над группами и индивидами к массовому мониторингу в поисках того, что можно назвать «разведданными», и БД это хорошо иллюстрируют. Практики слежки, раскрытые Э. Сноуденом, ясно показывают, что правительства разных стран (американское, британское, канадское)

и возможно другие агентства вовлечены в поразительно крупные масштабы мониторинга населения. Д. Лион вполне правомерно отмечает: «То, что пользователи по неведению раскрывают на таких платформах, как Facebook или Twitter, или при использовании своих телефонов, это ценные данные для «национальной безопасности» и полицейских целей. Но что еще более важно, с точки зрения Больших Данных, метаданные, касающиеся пользователей, собираются без их ведома, из простого использования этих аппаратов. Таким образом, существует по крайней мере три основных актора в этой драме: правительственные агентства, частные корпорации и, пусть и невольно, обычные пользователи» [8, р. 3].

Парадоксальное следствие современной ситуации расширения прозрачности повседневной жизни самого широкого круга населения состоит в том, что инстанции, осуществляющие наблюдение, остаются невидимыми и неизвестными для тех, кто находится под наблюдением.

Упреждающее слежение

Важно отметить, что появление БД сделало возможным новую форму наблюдения и слежения, которая осуществляется теперь с превентивными целями ввиду широко распространенной идеи «национальной безопасности». Превентивная, опережающая слежка теперь осуществляется *до*, а не после совершения преступления: «Большие Данные изменяют предшествующие полицейские или разведывательные мероприятия, которые условно были нацелены на подозреваемых или лица, представляющие интерес, а затем собирались данные о них. Теперь банк данных получается и собирается из различных источников до того как определяется полный спектр их актуальных и потенциальных использований и мобилизуемых алгоритмов и аналитики не только для понимания прошлой последовательности событий, но и для прогнозирования и вмешательства до действий, событий, процессов, устанавливаемых в цепочку» [8, р. 4].

Такой упреждающий подход является общим для всего спектра применения Больших Данных. Google Now, например, исполь-

зует этот метод для вывода огромного сцепления данных для того, чтобы предупреждать конкретных пользователей о вещах, которые могут серьезно повлиять на них, от задержки авиа-рейсов до раннего диагноза гриппа [8, р. 7]. Важно отметить наличие связи интересов между коммерческими и политическими интересами в использовании БД, при том, что политические интересы созвучны с идеологией неолиберализма: «Национальная безопасность – это бизнес-цель настолько, насколько и политическая цель» [8, р. 9].

Источники данных при этом могут быть различны: 1) направленный поиск оператора (например по системе видеонаблюдения, данных о владельце автомобиля и др.) 2) автоматический сбор данных, без вмешательства оператора (банковские транзакции, переговоры по мобильной связи), 3) «добровольные» данные, когда пользователь предоставляет свои данные в социальных медиа (хотя, как было отмечено, делает это не целенаправленно). Предсказаний человеческого поведения может быть несколько видов: а) *последовательное*, где одной из целей является помощь клиентам или пользователям выбрать то, что, скорее всего, будет выгодно для них, б) *преференциальное*, когда маркетологи пытаются угадать наши желания на базе наблюдения за нашим поведением, и в) *превентивное*, где есть явное намерение сократить чей-то спектр возможностей. По мнению Д. Лиона, в контексте справедливости и права именно последний вид поднимает фундаментальные вопросы приватности и надлежащей правовой процедуры. «Где правовые системы основаны на постфактумной системе штрафов или наказаний, поворот к системе, основанной на ориентированных в будущее профилактических мер, имеет огромное значение» [8, р. 5].

Интересно, что не взирая на выделенные проблемные стороны использования БД, фокусировка на негативных последствиях, способных привести к сокращению демократических свобод, к необратимой трансформации частной жизни в сторону большей уязвимости и роли информации в современном мире, крайне маргинализирована и не вызывает полноценного общественного интереса. Сказанное ставит вопрос с множеством разветвлений о том, в каком обществе мы хотим

жить? Будет ли это оруэлловское общество, где обычные люди никак не защищены от влияния непрозрачных организаций, «общество, отмеченное страхом и взаимной подозрительностью, где данные собираются без разбора и хранятся вечно в системах, которые никогда не забывают, делая прощение устаревшим, и создающих сильный страх, хотя тебе нечего скрывать?» [8, р. 11]. Пока не выработано ясных норм, позволяющих защищать конфиденциальность данных: «современное состояние Больших Данных для научного сообщества все еще Дикий Запад» [7]. В конечном счете этический поворот в литературе о БД становится более необходим как способ социальной критики.

Открытость

Новая ситуация, названная в литературе «Web 2.0» и означающая первичность интернет-активности самих пользователей в форме ведения блогов, социальных сетей, поддержки Википедии и т. п., усложняет вопрос о приватности данных и переводит его несколько в иную плоскость. Одни данные должны быть закрыты, а другие, наоборот, открыты. Как отмечает К. Лагозе: «Ожидания открытого доступа к информации, активного участия в производстве и комментировании знаний, а также интеграции социальной деятельности и познавательной деятельности являются сейчас нормой» [6, р. 7].

Распространение данных это общественное, публичное благо, соответственно и защита их конфиденциальности также должна полагаться на широкие кооперационные усилия по выработке приемлемых практик. «Действительно, технологические факторы, которые привели одновременно к взрывному росту в возможности предоставления таких данных, а также в способности злоумышленников нарушать конфиденциальность, находятся в зоне ответственности компьютерных ученых и социологов. Сообщество практиков может быть развито так, чтобы включить в него исследователей, которым будет поручено инициировать и воспитывать сообщество в стиле Web 2.0, которое сфокусируется на продвижении безопасных средств доступа к конфиденциальным данным» [7].

Кроме того, сегодня существует множество различных движений, выступающих за открытость данных, например, Германский Фонд Открытого Знания (OKF DE). Исследования С. Баак [2] показали, что активисты этого фонда демонстрируют как минимум три способа использования открытого исходного кода: во-первых, рассматривая данные как необходимое условие для формирования знаний, активисты трансформируют совместное использование исходного кода до включения совместного использования первичных данных (rawdata). Разделение первичных данных должно расколоть интерпретативную монополию правительств и позволит людям сделать свою собственную интерпретацию данных, касающихся социальных вопросов. Во-вторых, активисты соединяют эту идею с открытой и гибкой формой представительной демократии, применяя модель открытого участия к политическому участию в процессах принятия решений и активности в локальных сообществах. В-третьих, активисты признают, что необходимы посредники, чтобы сделать первичные данные доступными для общественности. Это, в частности, обуславливает их заинтересованность в трансформации журналистики, чтобы та стала посредником в этом смысле. В то же время они и сами пытаются выступать в качестве посредников и развивать гражданские технологии для воплощения своих идей в жизнь. Эти три направления имеют потенциал решить вопросы, связанные с опасностью того, что технологии БД подрывают демократическую свободу. Активисты нацелены на создание нового смысла легитимности производства и распределения коллективного знания в рамках демократической «датафицированной» (data fied) общественности. «При всех их недостатках, эти попытки являются актуальными и заслуживают внимания, поскольку они обеспечивают жизненно важную отправную точку, для обсуждения того, как мы можем противостоять угрозам Больших Данных и использовать потенциал этих новых технологий в целях, не подрывающих демократические ценности и свободу выбора тех, кем не в большой политике или крупном бизнесе» [2, р. 8].

В целом движения за открытые данные интересны тем, что они пересекаются с дву-

мя современными трендами, противоречащими друг другу: 1) «датафикация» (data fication) и 2) хакерство или культура открытого исходного кода. С одной стороны, практики и идеология активистов за открытые данные сосредоточены вокруг распределения и использования данных и, следовательно, привязаны к датафикации и вездесущей квантификации общественной жизни, для чего БД являются наиболее ярким отражением. С другой стороны, активисты за открытые данные применяют практики и ценности культуры открытого исходного кода к созданию и использованию новых данных. «Это связывает их с другими инициативами культуры открытого кода, как Open Access, Wikipedia, Wikileaks, Anonymous или Creative Commons... Отличная от датафикации, однако, культура открытого исходного кода связана с прозрачной и совместной формой управления» [2, р. 2]. Очевидно, что последняя стремится с использованием идеи справедливости (в форме принципов совместности, прозрачности и подконтрольности) гуманизировать условно объективный процесс всесторонней информатизации социокультурной реальности.

Пересмотр справедливости?

Ник Колдри призывает обратить внимание на потенциал одной из форм дискурса о БД монополизировать интерпретацию социокультурной реальности и тем самым скрыть существующие формы несправедливости: «...миф о Больших Данных рискует, устанавливая себя как здравый смысл, способ мышления о человеческой деятельности, что подрывает сами языки, посредством которых мы говорим о справедливости и несправедливости. Почему? Потому что версия социального знания мифа о Больших Данных не содержит никакого интерпретативного языка, из которого сравнения, как обстоят дела для разных людей, могут быть построены» [4, р. 11]. Если современная ситуация использования индивидуальных данных и других спорных форм организации социального порядка будет подаваться как неизбежная, объективная и неспецифицированная в вопросе ее бенефициаров и ответственных субъектов, то о ней нельзя будет говорить как о справедливой или неспра-

ведливой. «Такие несправедливости трудно назвать именно потому, что они связаны с концентрациями власти над ресурсами для именования. <...> Властные асимметрии, вовлеченные в «медийные центры», настолько вписаны в организацию современности и его пространства для требования справедливости, что трудно сейчас понять, как мы можем действовать без них; действительно, когда подлинные несправедливости происходят в результате деятельности СМИ (британские скандалы со взломом телефонов), то их трудно устранить. Но несправедливости, связанные с новыми мифами о нас и о Больших Данных могут иметь даже более серьезные последствия в долгосрочной перспективе» [4, р. 10].

В. Майер-Шенбергер и К. Кукьер считают необходимым обратить внимание на наши представления о справедливости из других оснований. «Мировоззрение, сотканное из понимания причин, теперь оспаривается доминированием корреляции. Обладание знанием, которое когда-то означало понимание прошлого, постепенно преобразовывается в способность прогнозировать будущее» [1, с. 198]. Такой пересмотр нужно сделать по двум причинам. Во-первых, ради того чтобы: «сохранить понятие «человеческого фактора» – свободы воли, согласно которой люди сами выбирают, как им действовать» [1, с. 182]. Если раньше человек оценивался по его поступкам, по тому что он сделал его судили морально и юридически, то в эпоху БД можно выявлять склонности человека и судить о человеке или судить человека, исходя из прогноза о том, что он сделает. Во-вторых, чтобы «включить меры по обеспечению безопасности человеческого фактора» [1, с. 182]. То есть, во-первых, нужно сохранить человеческую свободу, а, во-вторых, ее обезопасить. «Защита частной жизни потребует от лиц, имеющих дело с личными данными, большей ответственности за свою политику и действия. Нам предстоит пересмотреть свое представление о справедливости, чтобы гарантировать человеческое право на свободу действий (и, конечно, соблюдение ответственности за эти действия). Понадобятся новые учреждения и эксперты (так называемые «алгоритмисты»), чтобы интерпретировать сложные алгоритмы, на основе которых формируются выводы из

больших данных, и защищать интересы тех, кто может от этих выводов пострадать, например получить отказ в приеме на работу или хирургическом вмешательстве или не получить кредит из-за того, что о них «говорят» большие данные. Дело не в адаптации существующих правил, а в создании новых» [1, с. 178].

Авторы предлагают три стратегии, позволяющие справедливо и эффективно управлять информацией в новую эпоху БД: «смещение защиты конфиденциальности от индивидуального согласия в сторону подотчетности пользователей данных, закрепление приоритетности человеческого фактора над прогнозами, а также создание нового класса аудиторов больших данных (алгоритмистов) – могут служить основой эффективного и справедливого управления информацией в эпоху больших данных» [1, с. 192].

Цифровое неравенство

Тем не менее предложенные направления переосмысления справедливости не исчерпывают всех проблем, которые имеются сегодня в информационной области. В частности, уже цитированные выше Д. Бойд и К. Кроуфорд обращают внимание на то что ограниченный доступ к БД производит новые цифровые разделения. Это связано с тем, что исторически сбор и накопление данных был тяжелым, трудоемким и ресурсоемким процессом. Большая часть энтузиазма, окружающего БД сегодня, проистекает из того восприятия, что теперь имеется легкий доступ к большим объемам данных. Однако это не так. В настоящее время только крупные компании социальных медиа имеют доступ к действительно большим объемам социальных данных. При этом многие компании (IBM, Microsoft и др.), обладающие данными, не готовы делиться информацией со всеми желающими. Исследователи зачастую только по этой причине становятся сотрудниками таких компаний или устраивают туда своих студентов и аспирантов, чтобы иметь полноценный доступ к их базам данных. «Некоторые компании ограничивают доступ к своим данным полностью; другие продают привилегию доступа за плату; и другие предлагают небольшие наборы данных для университетских исследователей. Топовые и богатые

вузы могут покупать доступ к данным, и студенты ведущих университетов являются наиболее вероятными работниками, которых пригласят поработать в крупных социальных медиа-компаниях. Студенты из периферии с меньшей вероятностью могут получить эти приглашения и развить свои навыки. В результате различие между учеными будет значительно возрастать» [3, р. 673–674]. Таким образом, само пользование БД порождает и будет порождать неравенства в доступе к ним. «Когда вычислительные навыки позиционируются как самые ценные, возникают вопросы по поводу преимуществ в таком контексте. Это в своем роде устанавливает новые иерархии вокруг того, «кто может читать цифры», вместо того, чтобы признать, что и компьютерные ученые и социологи могут предложить ценный опыт» [3, р. 674].

Таким образом, существующая экосистема вокруг БД создает новый вид цифрового разрыва: на богатых и бедных в сфере БД. Л. Манович пишет в этой связи о трех «дата-классах» в «обществе Больших Данных»: «те, кто создает данные (как осознанно, так и благодаря «цифровому следу»), тех, кто имеет средства, чтобы собирать их, и те, кто имеют опыт анализа» [9, р. 470]. Последняя группа является самой маленькой, и одновременно самой привилегированной: «они определяют правила того, как Большие Данные будут использоваться, и кто будет в этом участвовать. В то время как институциональное неравенство может быть неизбежным следствием в научных кругах, они должны тем не менее рассматриваться и анализироваться. Они производят необъективность в возникающих данных и в способах исследований» [3, р. 675].

Можно выделить три стадии цифрового разделения: 1) доступ к технологии, 2) ее эффективное использование, 3) социальная интеграция и технологический эффект. Быстро сужающийся разрыв между индустриальными и развивающимися странами делают последние две стадии более существенными сегодня [5, р. 2]. Тем самым цифровое разделение и неравенство может иметь глобальное измерение – развивающиеся и неразвитые страны оказываются в менее выгодном положении и не всегда способны воспользоваться полезными возможностями БД. Н. Кшетри анализирует важные аспекты нашего по-

нимания факторов и ингибиторов диффузии эффективного использования БД в развивающихся странах, а также отношения к текущей природе интернационального цифрового разделения: «Равенство доступа является необходимым, но не достаточным условием, чтобы привести к цифровому равенству. Такие факторы, как наличие навыков и возможностей, социальные и культурные установки в отношении технологии, институциональная среда и социальные преобразования тесно связаны с возможностью использовать технологию эффективно и правильно» [5, р. 5].

Наконец, нужно отметить *дискриминационный потенциал* БД. Большие данные позволяют нам создавать обобщения человеческого поведения в прежде невозможных масштабах. Использование для работы небольшой выборки, извлеченной из больших данных, представляет собой важную эмпирическую установку для количественной социальной науки. Б. Веллес призывает не отказываться от возможности использовать Большие Данные как «малые» с целью избежать ошибок прошлого и уделить должное внимание социально-политическим меньшинствам, поскольку в стремлении производить эти обобщения, мы рискуем потерять из виду второстепенное, но немаловажное преимущество Больших Данных – обильную представленность меньшинств. «Когда женщины и меньшинства исключены из числа субъектов основного социального исследования, существует тенденция идентифицировать опыт большинства как «нормальный» и обсуждать опыт меньшинств в терминах того, насколько они отклоняются от этих норм» [13, р. 2]. Такая позиция представляется нам вполне обоснованной. «Вместо того, чтобы активно убирать эти голоса через выборки и чистку данных, или через их пассивное замалчивание путем статистического агрегирования, я решил принять возможность изучить статистические резко выделяющиеся значения, и я призываю своих коллег сделать то же самое. Делая Большие Данные небольшими, мы можем исправить исторические упущения и перекосы в социально-научных исследованиях и построить лучшее, большее и всеобъемлющее понимание человеческого поведения» [13, р. 2].

Таким образом, БД как новый социально-технический феномен наряду с впечатляющими новыми познавательными возможностями породил целый ряд фундаментальных морально-этических вопросов. «Эра Big Data только наступила, но уже сейчас важно, чтобы мы начали анализировать предпосылки, ценности и искажения этой новой волны исследований» [3, р. 675]. Одним из преломлений этих вопросов становится идея справедливости – центральная регулятивная общественная ценность, имеющая различные способы интерпретации. В этой статье были намечены некоторые из аспектов критики БД с позиции общего дискурса справедливости. В частности, персональные данные и их приватный характер, вступающий в конфликт с современными требованиями национальной безопасности и практиками превентивного слежения; проблема открытости информации; неравный доступ к информации и БД, порождающий новый тип социальной несправедливости – цифровое неравенство, выражающееся в расслоении из-за неравенства в доступе к информационным технологиям; проблема дискриминации социальных меньшинств в использовании БД. Мы надеемся, что научное сообщество сможет предложить и свои варианты как решения названных вопросов, так и вычленения других, не менее важных, аспектов проблемы справедливости в информационной области.

ПРИМЕЧАНИЕ

¹ Статья написана при поддержке гранта РФФИ в рамках научно-исследовательского проекта 15-06-02758А «Реляционное объединение данных разнородных социологических исследований в единый массив и интеллектуальный анализ с применением технологий машинного обучения».

СПИСОК ЛИТЕРАТУРЫ

1. Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / В. Майер-Шенбергер, К. Кукер ; пер. с англ. И. Гайдюк. – М. : Манн, Иванов и Фербер, 2014. – 240 с.
2. Baack, S. Datafication and empowerment: How the open data movement re-articulates notions

of democracy, participation, and journalism / S. Baack. – Big Data & Society, 2015. – pp. 1–11. – DOI: 10.1177/2053951715594634.

3. Boyd, D. Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon / D. Boyd, K. Crawford // Information, Communication & Society. – 2012. – Vol. 15, № 5. – pp. 662–679.

4. Couldry, N. A necessary disenchantment: myth, agency and injustice in a digital world / N. Couldry // The Sociological Review. – 2014. – Vol. 62, № 4. – pp. 880–897. – DOI: 10.1111/1467-954X.12158.

5. Kshetri, N. The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns / N. Kshetri. – Big Data & Society, 2014. – pp. 1–20. – DOI: 10.1177/2053951714564227.

6. Lagoze, C. Big Data, data integrity, and the fracturing of the control zone / C. Lagoze. – Big Data & Society, 2014. – pp. 1–11. – DOI: 10.1177/2053951714558281.

7. Lane, J. Protecting Privacy and Confidentiality in an Era of Big Data Access / J. Lane. – Chance, 2012. – Vol. 256, № 4, pp. 39–41. – Electronic text data. – Mode of access: <http://chance.amstat.org/2012/11/privacy25-4/>. – Title from screen.

8. Lyon, D. Surveillance, Snowden, and Big Data: Capacities, consequences, critique / D. Lyon. – Big Data & Society, 2014, pp. – 1–13. – DOI: 10.1177/2053951714541861.

9. Manovich, L. Trending: the promises and the challenges of big social data / L. Manovich // Debates in the Digital Humanities. – The University of Minnesota Press, 2011. – pp. 460–475.

10. Ovadia, S. The Role of Big Data in the Social Sciences / S. Ovadia. – Behavioral & Social Sciences Librarian, 2013. – Vol. 32, №. 2. – pp. 130–134. – DOI: 10.1080/01639269.2013.787274.

11. Peacock, S. E. How web tracking changes user agency in the age of Big Data: The used user / S. E. Peacock. – Big Data & Society, 2014. – pp. 1–11. – DOI: 10.1177/2053951714564228.

12. Vaidhyanathan, S. Bullock C. Knowledge and Dignity in the Era of “Big Data” / S. Vaidhyanathan, C. Bullock // The Serials Librarian: From the Printed Page to the Digital Age. – 2014. – Vol. 66, № 1–4. – pp. 49–64. – DOI: 10.1080/0361526X.2014.879805.

13. Welles, B. F. On minorities and outliers: The case for making Big Data small / B. F. Welles. – Big Data & Society, 2014. – pp. 1–2. – DOI: 10.1177/2053951714540613.

REFERENCES

1. Mayer-Schönberger V., Cukier K. *Bolshie dannye. Revolyuciya, kotoray izmenit to, kak my*

zhivem, rabotaem i myslim [Big Data: A Revolution that Will Transform How We Live, Work, and Think]. Moscow, Mann, Ivanov i Ferber Publ., 2014. 240 p.

2. Baack S. Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2015, pp. 1-11. DOI: 10.1177/2053951715594634.

3. Boyd D., Crawford K. Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 2012, vol. 15, no. 5, pp. 662-679.

4. Couldry N. A necessary disenchantment: myth, agency and injustice in a digital world. *The Sociological Review*, 2014, vol. 62, no. 4, pp. 880-897. DOI: 10.1111/1467-954X.12158.

5. Kshetri N. The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, 2014, pp. 1-20. DOI: 10.1177/2053951714564227.

6. Lagoze C. Big Data, data integrity, and the fracturing of the control zone. *Big Data & Society*, 2014, pp. 1-11. DOI: 10.1177/2053951714558281.

7. Lane J. Protecting Privacy and Confidentiality in an Era of Big Data Access. *Chance*, 2012, vol. 256,

no. 4, pp. 39-41. Available at: <http://chance.amstat.org/2012/11/privacy25-4/>. (accessed August 5, 2015).

8. Lyon D. Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 2014, pp. 1-13. DOI: 10.1177/2053951714541861.

9. Manovich L. Trending: the promises and the challenges of big social data. *Debates in the Digital Humanities*. The University of Minnesota Press, Minneapolis, MN, 2011. pp. 460-475.

10. Ovadia S. The Role of Big Data in the Social Sciences. *Behavioral & Social Sciences Librarian*, 2013, vol. 32, no. 2, pp. 130-134. DOI: 10.1080/01639269.2013.787274.

11. Peacock S.E. How web tracking changes user agency in the age of Big Data: The used user. *Big Data & Society*, 2014, pp. 1-11. DOI: 10.1177/2053951714564228.

12. Vaidhyanathan S. & Bullock C. Knowledge and Dignity in the Era of "Big Data". *The Serials Librarian: From the Printed Page to the Digital Age*, 2014, vol. 66, no. 1-4, pp. 49-64. DOI: 10.1080/0361526X.2014.879805.

13. Welles B.F. On minorities and outliers: The case for making Big Data small. *Big Data & Society*, 2014, pp. 1-2. DOI: 10.1177/2053951714540613.

JUSTICE IN THE INFORMATION AREA: ETHICAL ASPECTS OF BIG DATA

Evgeniy Vladimirovich Karchagin

Candidate of Philosophical Sciences, Associate Professor,
Department of Philosophy, Sociology and Psychology,
Volgograd State University of Architecture and Civil Engineering
evgenkar@yandex.ru
Akademicheskaya St., 1, 400074 Volgograd, Russian Federation

Abstract. The article deals with the problem of justice in relation to the current situation, due to the emergence of a new source and method of processing information - Big Data. The author stands out several sources of the need to rethink the justice in information area: personal information and its private character coming into conflict with modern security requirements and practices of proactive monitoring; public movement for open data; unequal access to Big Data, which generates a new type of social injustice – the digital inequality, which is expressed in disparity because of inequalities in access to information technology; the issue of discrimination and the use of Big Data. The paradoxical consequence of the current situation of expansion of transparency of daily life of population is that the authorities responsible for monitoring remain invisible and unknown to those who are supervised. Movements for open data are interesting because they intersect with two modern trends: 1) datafication; 2) hacking and open source culture. Existing ecosystem around Big Data creates a new kind of digital division: the rich and the poor in Big Data sphere. Large data can be used as “small” in order to avoid mistakes of the past and give consideration to the minorities, because in an effort to make generalizations we risk to lose sight of the secondary, but important advantage of Big Data is the abundant representation of minorities.

Key words: Big Data, justice, ethics, information area of justice, privacy and publicity of data, digital inequality.