

УДК 159.9

АНАЛИЗ БОЛЬШИХ ДАННЫХ В ПСИХОЛОГИИ И СОЦИОГУМАНИТАРНЫХ НАУКАХ: ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ¹

© 2019 г. Т. А. Нестик*, А. Л. Журавлев**

**Доктор психологических наук, профессор РАН,
зав. лабораторией социальной и экономической психологии.*

E-mail: nestik@gmail.com

***Академик РАН, научный руководитель ИП РАН.*

E-mail: alzhuravlev2018@yandex.ru

Поступила 29.08.2019

Аннотация. Рассматриваются результаты сбора и учета больших данных, характеризующих реальное поведение людей, а также последствия влияния больших данных на общество, социальные науки и психологическое знание. Формулируется определение понятия “большие данные”. Анализируются психологические исследования с использованием больших данных: их вторичный анализ; краудсорсинговые проекты и сопоставление “цифровых следов” пользователей социальных сетей с их ответами на стандартизированные методики; исследования психологических феноменов с применением “цифровых следов” и алгоритмов машинного обучения без проведения опросов; естественные эксперименты и сбор данных с помощью носимых электронных устройств. Отдельное внимание уделяется психологическому профилированию личности по “цифровым следам”. Предложены перспективные направления исследований динамики социально-психологических феноменов на личностном, межличностном, групповом и макросоциальном уровнях с использованием автоматизированного анализа естественного языка и изображений. Намечены психологические и социальные риски, связанные с внедрением в повседневную жизнь общества технологий анализа больших объемов данных.

Ключевые слова: большие данные, “цифровые следы” личности, визуальные данные, социальные сети, групповая динамика, макропсихологические показатели, риски.

DOI: 10.31857/S020595920007306-2

АНАЛИЗ “ЦИФРОВЫХ СЛЕДОВ” ЧЕЛОВЕКА: НОВЫЕ ВОЗМОЖНОСТИ ДЛЯ СОЦИОГУМАНИТАРНЫХ НАУК

Формирование “интернета всего”, накопление “цифровых следов” индивидуальной и совместной жизнедеятельности в сети людей сегодня ставит перед социогуманитарными науками задачу нового описания всей социальной действительности на основе данных о реальном поведении людей [2; 12]. Не только в социальных науках, но и в гуманитарном знании — истории, археологии, культурной антропологии, лингвистике, теории литературы и искусствоведении — анализ больших данных позволяет обнаружить закономерности, ранее невидимые глазу исследователя [13].

Ключевые отличия больших данных от традиционных “малых данных” в психологии и социогуманитарных науках сводятся к тому, что они огромны по объему и, как правило, не могут быть обработаны на персональном компьютере; они меняются в режиме реального времени и часто требуют изменения алгоритмов обработки; они разнообразны по составу, включают государственную статистику, тексты и метаданные, одним словом, требуют агрегирования совершенно разных по структуре массивов информации, что затрудняет использование универсального алгоритма их получения и анализа.

Для сбора “цифровых следов” используются программы-краулеры, автоматически отыскивающие и скачивающие данные по заданным критериям, а при изучении социальных сетей — приложения, которые подключаются через интерфейс *Twitter*, *Facebook*, *ВКонтакте* и других сетей

¹ Исследование выполнено по гранту РНФ № 18-18-00439.

(*API — Application Programming Interface*), в том числе различные автоматизированные опросы, представляющие мгновенную обратную связь респондентам. В психологических исследованиях для анализа больших данных используют в основном статистические пакеты *R* и, реже, *Python*. Используется логистический регрессионный анализ, “деревья” решений, иерархический кластерный анализ и др. Хотя некоторые методы анализа, предполагающие разбиение всей совокупности данных на независимые выборки с последующим метаанализом полученных результатов, допускают использование также и привычных инструментов, таких как *SPSS*. Для определения эмоциональной окрашенности текстов и выделения лингвистических маркеров психологических характеристик используются специальные программы, например *Linguistic Inquiry and Word Count* или *Text Analysis, Crawling and Interpretation Tool*, а также другие психолингвистические анализаторы [14].

Вместе с распространением технологий *Big Data* и “интернета всего” меняются скорость, масштабы и методы сбора данных в психологических исследованиях [30]. Радикально меняется относительная доступность для изучения четырех *типов данных*, выделенных Р. Кеттелом, а именно: *L*-данных (объективных фактов о жизни и поведении) и *T*-данных (результаты тестов) становится гораздо больше, чем *Q*-данных (опросников, основанных на самоотчетах респондентов) и *O*-данных (оценок человека другими знающими его людьми). Семантический анализ *SMS*, интернет-запросов и постов в сетях, а также обработка “цифровых следов”, которые человек оставляет в реальном мире, соприкасаясь с сетью электронных устройств, не говоря уже о показаниях сенсора движения в смартфоне и триангуляции его положения по системам *GPS*, — все это уже сегодня позволяет оценивать некоторые психологические характеристики человека более точно, чем с помощью стандартизированных психологических опросников [32]. В мире интернета вещей вся жизнь человека становится цифровой, доступной для наблюдений и диагностики — со всеми вытекающими отсюда позитивными и негативными последствиями. При этом появляется возможность не только судить о прошлом человека, но и предсказывать его поведение. В связи с тем что фокус внимания исследователей смещается с объяснения поведения на его прогнозирование, некоторые эксперты говорят о возможности преодоления кризиса воспроизводимости результатов психологических исследований с помощью больших данных [28].

Следует признать, что использование больших данных в научных психологических исследованиях затруднено целым рядом *барьеров*.

Во-первых, большой объем данных создает иллюзию их полноты, тогда как на самом деле они не обеспечивают репрезентативности и регулярности охвата. Например, они могут не учитывать характеристики людей, отсутствующих в социальных сетях; при лингвистическом анализе алгоритм может не читать какую-то часть текстов, тогда как мы считаем корпус полным [7].

Во-вторых, алгоритмы соотнесения “цифровых следов” с социально-демографическими и психологическими характеристиками, как правило, являются коммерческой тайной, что затрудняет не только воспроизводимость полученных результатов, но и стандартизацию процедур обработки данных. Разработка таких алгоритмов финансируется компаниями для решения узконаправленных маркетинговых или управленческих задач, которые не требуют комплексной психологической диагностики, соблюдения строгих критериев надежности и научно обоснованной интерпретации. В итоге создаются выборки, учитывающие в основном демографические показатели или построенные на популярных типологиях личности, а у программистов возникает иллюзия того, что нейросеть сама найдет все необходимые маркеры, поэтому психологов привлекать к этому не следует.

В-третьих, платформы, обрабатывающие большие данные, периодически изменяют алгоритмы расчета, что приводит к ошибкам в моделях прогнозирования, а также при сборе и обработке таких разнородных данных нередко имеют место смещение выборки, ошибки измерений, ложные корреляции и т.п.

В-четвертых, формулировка задач и интерпретация аналитиком результатов анализа больших данных не свободны от множества когнитивных искажений, некоторые из них становятся даже более вероятными, чем при работе с малыми данными. Безусловно, анализ больших данных расширяет возможности количественных методов проверки психологических гипотез. Вместе с тем обращение психологов к *Big Data* повышает востребованность качественных методов анализа и теоретического обоснования полученных результатов. Не случайно перспективы использования больших данных в социогуманитарных науках связывают с необходимостью дополнять их качественными, например этнографическими, исследованиями.

В-пятых, большие данные не бывают абсолютно “сырыми”, они хранятся с определенной целью,

включены в систему политических и экономических отношений, содержат преднамеренные и непреднамеренные пропуски. Иными словами, надежные психометрические алгоритмы не защищают исследователя от искаженных данных [28].

В-шестых, еще одним ограничением являются этические и юридические проблемы, связанные с использованием больших данных в психологических исследованиях. Соблюдение принципа информированного согласия в таких исследованиях не всегда возможно, а традиционное обезличивание данных не исключает возможности идентификации личности по совокупности “цифровых следов”, что заставляет в ряде случаев прибегать к шифрованию и уникальным ключам, чтобы ограничить использование данных. Ограничения, накладываемые этическими комитетами социальных сетей на доступ исследователей к данным, не всегда действенны. Так, например, в 2018 г. получил известность и продолжение скандал, связанный с использованием больших данных во время президентских выборов в США (2016 г.), в чем была уличена компания “Cambridge Analytica”, специализирующаяся на психологическом профилировании по “цифровым следам”. Появление владельцев данных (интернет-гигантов, социальных сетей, банков, страховых компаний, платежных систем и т.п.), которые, возможно, знают своих пользователей лучше, чем они сами, создает целый ряд этических и социальных проблем, не говоря о фобиях слежения в интернете, которыми тоже придется заниматься психологам.

Наконец еще один барьер связан с тем, что анализ больших данных — задача, посильная только междисциплинарным командам, состоящим из психологов, лингвистов, программистов и математиков. На наш взгляд, освоение больших данных психологами затруднено не только необходимостью обучения дополнительным навыкам [17], но и относительной замкнутостью исследовательского психологического сообщества.

ПСИХОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ДАННЫХ

На наш взгляд, по дизайну исследования работы психологов с использованием больших данных можно разделить на четыре *основные группы*. Во-первых, исследования, опирающиеся на вторичный анализ уже собранных и структурированных данных. Во-вторых, краудсорсинговые проекты, позволяющие собрать очень большие выборки участников исследования [23], а также сопоставить

“цифровые следы” пользователей социальных сетей с их ответами на стандартизированные опросники [8]. В-третьих, это исследования психологических феноменов, использующие “цифровые следы” и алгоритмы машинного обучения, которые позволяют анализировать неструктурированные текстовые и визуальные данные без проведения опросов (см., например, [18]). В-четвертых, исследования, в которых сбор данных организован как естественный эксперимент [41] или, например, с помощью очков расширенной реальности, смартфонов и планшетов.

Благодаря социальным сетям и инструментам краудсорсинга, таким как “Механический турок” компании “Amazon”, психологам стало легче собирать большие выборки. К давно открытым многолетним базам социологических опросов, таким как “World Values Survey”, “International Social Survey”, “European Social Survey” и “Eurobarometer” (см.: <http://www.re3data.org>), добавляются новые, собственно психологические (например, “Journal of Open Psychology Data”, “Cognitive Economics Project”, “International Teston Risk Attitude” и др.). Например, в “Experience Project” — социальной сети, существовавшей с 2007 по 2016 г., — хранится 66 млн жизненных историй и описаний различных эмоциональных переживаний, составленных ее участниками. “Speech Home Project” содержит 200 тыс. часов записей речи детей до 3 лет. Объем данных, накопленных когнитивными психологами и нейробиологами в рамках проекта “Human Connectome Project”, приближается к петабайту. На сайте *myPersonality.org* было собрано и до недавнего времени находилось в открытом доступе более 6 млн результатов тестов (в том числе 3.1 млн результатов по тесту Big5) вместе с 4 млн личных страниц в Facebook. Доступность психологических данных растет вместе с переходом психологии к модели “Открытая наука”. Все больше журналов требуют размещения в открытом доступе баз с сырыми данными, а такие как “Cortex”, “Attention Perception & Psychophysics”, “AIMS Neuroscience” и “Experimental Psychology” — даже предварительной публикации и одобрения дизайна исследования до проведения эксперимента.

Наибольшую известность получили исследования поведения пользователей сетевых сообществ [16] и социальных сетей [1], а также психологическое профилирование личности по “цифровым следам” на основе модели “Большой пятерки”, быстро взятое на вооружение в политических, маркетинговых и корпоративных проектах [3; 8].

Технологии анализа “цифровых следов” как маркеров психологических характеристик были

разработаны при изучении сообщений блогеров в *Twitter* и опирались преимущественно на компьютерную лингвистику. Прорывом в этой области стало сопоставление личных страниц в *Facebook* с ответами их владельцев на стандартизированные психологические опросники, осуществленное в рамках проекта *MyPersonality.org*, работы Психометрического центра Кембриджского университета и проекта “*World Well-Being Project*” Центра позитивной психологии Университета Пенсильвании. Начав в 2007 г. с рассылки приглашения 150 друзьям в социальной сети, М. Косинский и Д. Стилвелл за четыре года методом снежного кома собрали более 6 млн участников исследования [30]. Уже к 2013 г. им удалось обеспечить прогностическую валидность отметок “like” для шкал пятифакторной модели личности *Big Five* до 0.43, а при оценке политических убеждений и социально-демографических характеристик коэффициенты составляли от 0.7 до 0.9 [29]. Проведенный в 2017 г. метаанализ исследований, посвященных связи “цифровых следов” с “Большой пятеркой”, показывает, что их предсказательная сила колеблется от 0.29 для доброжелательности до 0.40 для экстраверсии, т.е. не уступает стандартизированным опросникам [19]. Разработаны алгоритмы, позволяющие судить об уровне интеллекта, удовлетворенности жизнью, склонности к самораскрытию и самомониторингу, ценностных ориентациях личности и характеристиках временной перспективы.

Для психологического профилирования личности используются не только лингвистические маркеры и сетевой анализ, но и фотографии, размещенные на страницах *Instagram* и *Facebook*, видеозаписи, аудиозаписи речи, видеоблоги и датчики движения в смартфонах. Для получения цифровых маркеров психологических характеристик исследователи, как правило, совмещают: 1) компьютерный анализ текста при помощи систем автоматизированного распознавания речи и невербального поведения (чаще всего *LIWC*); 2) самоотчеты испытуемых и экспертные оценки, в том числе кодирование данных экспертами; 3) машинное обучение. Растущее количество видеокамер в городах и совершенствование алгоритмов автоматизированного распознавания психологических состояний по поведенческим индикаторам превращают анализ потокового видео в наиболее перспективное направление психометрических исследований [28].

Сетевой и автоматизированный анализ естественного языка позволяет изучать психологические феномены без проведения опросов. В качестве иллюстрации возможностей и ограничений

такого подхода можно привести исследования распространения лжи в социальных сетях. Например, команда специалистов из Массачусетского университета проанализировала 120 тыс. новостей, которыми 4.5 млн раз поделились 3 млн пользователей *Twitter* в 2006–2017 гг. [44]. Сообщения были разделены на правдивые и ложные с опорой на 6 независимых фактчекинговых организаций. После выявления и устранения сообщений, сгенерированных ботами, оказалось, что ложные новости отличаются большей новизной, распространяются в 6 раз быстрее и в 10 раз шире, чем правда. Исследование показало, что в основе распространения дезинформации в сети лежат психологические закономерности массового сознания, а не только злой умысел “фабрик троллей”. Ложные новости были связаны с лингвистическими маркерами удивления, страха и отвращения в перепостах, а правдивые новости — с предвосхищением, досадой, радостью и доверием. При этом авторы исследования признают необходимость проведения интервью и экспериментов для интерпретации полученных данных. Еще одним примером исследования, проведенного без использования опросов на этот раз в области клинической психологии, стало определение депрессии и склонности к суициду по “цифровым следам”. Исследователи отобрали три группы интернет-форумов, две из которых включали сообщества людей, проходящих лечение депрессивных расстройств, а также сообщества, для которых характерна идеализация смерти. Третья, контрольная группа была представлена нейтральными форумами, посвященными женской и мужской тематике, карьере и профессиональному развитию. Автоматизированный анализ текстов показал, что на форумах первых двух групп в 3 раза чаще используется абсолютистская лексика, т.е. такие слова, как “всегда”, “никогда”, “совсем” и т.п. [18].

Известным примером экспериментального психологического исследования, основанного на сборе больших данных, стал проект, в котором зрительное восприятие изучалось с помощью игры в мобильном приложении. Пользователи должны были “просвечивать” сканером багаж пассажиров в аэропорту, что позволило зафиксировать более миллиарда зрительных реакций. При этом исследователям не только удалось изучить, как влияют отвлекающие внимание раздражители на восприятие редко предъявляемых стимулов, но и динамику процесса тренировки внимания [34]. Как правило, при проведении таких экспериментов исследователи не раскрывают свои цели участникам. Например, в январе 2012 г. исследователи из Корнельского университета провели

эксперимент среди 689 003 пользователей сети *Facebook*. Для изучения эмоционального заражения они манипулировали доступностью контента пользователям: те, кому были более доступны негативные сообщения друзей, сами чаще писали негативно окрашенные тексты [31]. Поскольку интернет-пользователи не всегда в полной мере отдают себе отчет в том, что являются участниками эксперимента, использование этого метода остается дискуссионным по этическим соображениям.

Проведенный Институтом психологии РАН экспертный опрос, посвященный будущему психологии, свидетельствует о том, что революция больших данных долгое время оставалась практически не замеченной в российском исследовательском сообществе [4]. Тем не менее специально проведенный анализ базы *Web of Science* показал, что за последние 5 лет число психологических исследований с использованием *Big Data* выросло в 20 раз [6]. Сегодня большие данные все более активно используются не только в кибер- и медиапсихологии, но и в психиатрии, исследованиях памяти, внимания и зрительного восприятия [26], организационной и экономической психологии [39]. Благодаря новым открытым базам данных расширяются возможности для оценки эффективности тех или иных методов лечения в психиатрии [43], совместный анализ больших данных в психогенетических исследованиях проливает свет на новые закономерности [27]. Анализ социальных сетей позволяет выявлять цифровые маркеры суицидальных состояний, шизофрении, тревожных и депрессивных расстройств [36]. В организационной психологии на основе данных о реальном поведении членов малой группы открывается возможность не только для разработки новых методов оценки кандидатов при приеме на работу, но и для исследования того, как групповые феномены (внутригрупповое доверие, сплоченность, групповые представления и ментальные модели) влияют на совместную деятельность [25].

В настоящее время оценку индивидуальных психологических характеристик по “цифровым следам” можно в целом считать решенной научной задачей, что не отменяет необходимости уточнения соответствующих моделей для русскоязычного кластера интернета [9]. На наш взгляд, фундаментальный прорыв необходим и может быть достигнут в другой области — при прогнозировании *динамики* межличностных и групповых психологических и макропсихологических (характерных для общества) феноменов на основе “цифровых следов” реального поведения.

ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ СОЦИАЛЬНО-ПСИХОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ НА ОСНОВЕ БОЛЬШИХ ДАННЫХ

В этой части статьи хотелось бы наметить несколько *перспективных направлений исследования* динамики социально-психологических процессов с использованием анализа больших данных.

Во-первых, накапливающиеся на протяжении жизни человека “цифровые следы” становятся важным источником информации о динамике социально-психологических характеристик личности. На личностном уровне социально-психологического анализа лингвистические маркеры постов в социальной сети, тематика групп, на которые подписан пользователь, а также граф его контактов могут указать на изменения в характеристиках социальной идентичности, ценностных ориентациях, а также в социальных ролях в сетевых сообществах. Например, с “цифровыми следами” могут быть сопоставлены выявленные на основе самоотчетов интернет-пользователей роли в сетевом сообществе, такие как “творец”, “защитник”, “троль”, “наставник”, “посредник”, “наблюдатель”, “собеседник”, “актер”, “манипулятор”, “друг” [15]. Появляется возможность для уточнения моделей многих социально-психологических процессов: формирования ценностных ориентаций, вторичной социализации, социального сравнения и относительной депривации, формирования социального капитала личности и др.

Анализ “цифровых следов” открывает путь к решению задачи, трудно реализуемой через самоотчеты респондентов, — определению связи динамики социально-психологических характеристик личности с изменениями в макропсихологических характеристиках общества (о макропсихологии см.: [5; 10]). Такого рода работы уже появились. Например, исследование, проведенное среди 380 тыс. жителей Англии и Уэльса, показало, что в тех регионах страны, которые во время индустриальной революции пережили подъем и затем спад угольной промышленности, до сих пор сохраняются более низкие показатели сознательности и удовлетворенности жизнью, но более высокие показатели невротизма, тревоги и депрессии [37].

Во-вторых, относительно мало разработанной, но все более актуальной становится проблема прогнозирования динамики межличностных процессов по “цифровым следам”. Внимание исследователей до сих пор было сосредоточено в основном на изучении личностных характеристик интернет-пользователей. Между тем на межличностном

уровне и уровне малых групп семантический анализ текстов сообщений, мониторинг структуры социальной сети и интенсивности обмена сообщениями в мессенджерах делают возможным оценку и прогнозирование стадий развития межличностных отношений. Со временем можно ожидать появления облачных сервисов, которые будут предлагать семьям и парам, связанным романтическими отношениями, не только оценку их совместимости, но и прогноз потенциальных конфликтов, развитие эмоционального интеллекта и сензитивности к состояниям друг друга на основе мониторинга их сетевой активности. В последние годы появилось множество сайтов и мобильных приложений, обещающих своим пользователям подбор потенциальных партнеров на основе больших данных (*eHarmony, Hinge, Who's Here, Skout, Tinder* и др.). При этом эксперименты показывают, что сообщение пользователям об их совместимости действует как самосбывающееся пророчество вне зависимости от того, насколько они в действительности совместимы [40]. Согласно исследованию использования больших данных в сфере романтических отношений, проведенному *eHarmony* совместно с Лондонским имперским колледжем, в ближайшие 10 лет для оценки совместимости и изучения межличностных отношений будут все больше использоваться “умные вещи” — от носимых гаджетов до умных домов [24]. Эта динамично развивающаяся область прикладных разработок, часто сомнительных с точки зрения обоснованности исходных предположений и интерпретации данных, остается вне зоны внимания серьезных академических исследователей.

В-третьих, анализ “цифровых следов”, накопленных проектными командами и организациями в ходе совместной деятельности, открывает путь к изучению групповой динамики и влияния групповых феноменов на результаты совместной деятельности. Первые шаги в этом направлении уже делают разработчики корпоративного программного обеспечения. Так, например, *Microsoft Delve* строит графы связей между сотрудниками и анализирует содержание всех коммуникаций в интегрированной среде *Microsoft*, в том числе посредством мессенджера *Yammer*; сервис *Trust Vault* оценивает структуру социальной сети, лидерское поведение и уровень внутригруппового доверия на основе анализа переписки сотрудников; *Syndio* позволяет проводить опросы среди сотрудников и сопоставлять их данные с сетевой активностью, измеряя таким образом отношение к работодателю. А приложения *Mood Metrix* и *Attitude Heat map*, разработанные компанией *Keen Corp* в партнерстве с *Omniscien Technologies*, на основе психолингвистического анализа

электронной почты и мессенджеров оценивают эмоциональное состояние коллектива и вовлеченность сотрудников. Динамику эмоционального состояния можно ежедневно отслеживать на уровне всего коллектива, отдельного департамента и даже конкретной команды, состоящей минимум из 10 участников. Разработчики программы утверждают, что личная информация полностью защищена: данные анонимизируются, тексты сообщений кодируются и обрабатываются в агрегированном виде. Сопоставление “цифровых следов” в корпоративных мессенджерах и социальных сетях с результативностью проектных команд, экспертными оценками руководителей и данными опросов сотрудников позволяет разработать алгоритм, на основе которого можно будет не только оценивать социально-психологический климат в коллективе, но и прогнозировать успешность проектных команд.

Анализ текстов поможет глубже понять целый ряд групповых процессов, находящихся в фокусе внимания современной психологии малых групп и организационной психологии. Как формируется и меняется психологический контракт в малых группах и организациях? Как изменяется транзакционная память команды при частичном изменении состава участников? Каковы стадии формирования групповой идентичности, внутригруппового доверия, разделяемых членами группы ментальных моделей и ситуационной осведомленности? Какова последовательность стадий принятия организационных изменений в разных корпоративных культурах? Каковы факторы эффективности лидерского поведения в виртуальных, распределенных командах, участники которых работают удаленно? Совмещение опросов, анализа дискурса и кейс-стади с анализом больших данных позволяет прояснить закономерности групповой рефлексии и ее влияния на эффективность совместной деятельности.

В-четвертых, большие данные дают возможность для изучения групповых процессов, развивающихся одновременно в сетевом и реальном пространствах. Анализ меток геолокации в сообщениях пользователей социальных сетей открывает новые возможности для исследователей в сфере социальной психологии города, психологии криминального поведения, а также изучения межнациональной напряженности и агрессии. Появляется возможность строить географические карты, отображающие распространение нетерпимости в локальных сообществах. Становится возможным изучение динамики межгрупповых отношений с учетом региональных особенностей

и микроконтекста социальной ситуации. Геометрические сообщения, триангуляция звонков и точек подключения к сети носимых гаджетов, датчики движения в смартфонах — все это со временем расширяет возможности не только для изучения массовых феноменов, в том числе социальных движений, изменения поведения толпы под влиянием поступающих одновременно реальных и виртуальных сообщений разного психологического содержания, но и для изучения воздействия различных условий проживания в городе на психологическое благополучие горожан, различные психологические типы отношения жителей к новым транспортным и архитектурным технологиям “умного города”.

В-пятых, чрезвычайно перспективны цифровые исследования динамики психологических процессов в больших социальных группах, в том числе изучение изменений макропсихологических характеристик отдельных регионов и общества в целом. Сопоставление данных психолингвистического анализа текстов интернет-пользователей, административной статистики и результатов репрезентативных социологических опросов за большие периоды времени позволяет разработать объяснительные и прогностические модели динамики макропсихологического состояния отдельных регионов, а также давать прогноз влияния макропсихологических показателей на социально-экономическую и политическую ситуацию. Сочетание анализа цифровых следов, дискурса и социально-психологических онлайн-опросов расширяет возможности для изучения сложных психологических феноменов массового сознания: коллективных переживаний, в том числе мировоззренческих чувств, социальных представлений, коллективной памяти, коллективного образа будущего. До сих пор исследователи были сосредоточены в основном на динамике относительно простых групповых феноменов в социальных сетях: эмоциональном заражении и распространении мемов (см., например, [42]), информационных каскадах в связи с теми или иными обсуждаемыми событиями (см., например, [21]), групповой поляризации и возникновения “эхо-камер” [20], распространении слухов и ложной информации [11], тогда как исследования более сложных явлений, таких как коллективная память и социальные представления, относительно редки [22]. Между тем автоматизированный анализ текстов пользователей форумов и социальных сетей, а также публикаций в СМИ приближает нас к решению более сложных научных проблем: например, какие социально-психологические факторы влияют на динамику доверия к социальным институтам; каким образом содержание коллективной памяти влияет на социальные представления о будущем,

какие групповые социально-психологические характеристики влияют на динамику вовлечения жителей региона в использование технологических инноваций и на развитие производительных сил регионов страны.

Особое внимание необходимо обратить не только на механизмы распространения тревожных состояний и коллективных страхов, но и на социально-психологические закономерности распространения позитивной информации, позитивной оценки коллективного прошлого, формирования позитивной гражданской идентичности и ответственности за происходящее в стране. Например, “цифровые следы” позволяют изучать динамику коллективных переживаний во время праздников и траура, формирование генерализованной надежды, феномены “большого оптимизма” и “защитного пессимизма”, психологические функции коллективного образа ближайшего, среднесрочного и отдаленного будущего. Одним из прикладных применений этой работы может быть прогнозирование психологического воздействия “мнемонических войн” и информационных кампаний, направленных на изменение отношения общества к своему коллективному прошлому и будущему. В результате таких междисциплинарных исследований может быть разработана модель формирования групповой жизнеспособности в цифровой среде, а также устойчивости крупных социальных групп к информационным угрозам в условиях растущей неопределенности.

Большие данные не только позволяют заглянуть в сегодняшнее состояние российского общества, но и проливают свет на феномены исторической психологии. Использование исторических больших данных в психологических исследованиях имеет свои недостатки и ограничения (см. критический анализ [38]). Тем не менее их можно использовать для изучения содержания коллективной памяти и ее динамики, а также для изучения изменений коллективного образа будущего.

Большой интерес представляют в этой связи новые возможности для анализа визуальных сообщений. Сочетание автоматизированного распознавания изображений и экспертных оценок позволяет исследовать динамику социальных представлений и сложных коллективных эмоциональных переживаний, оценивать различия в визуальной репрезентации сложных явлений представителями разных социальных групп, жителями разных регионов. Методологические основания такого анализа для этого уже разработаны в рамках визуальной социологии, а также социальной психологии изображений [33; 35].

ЗАКЛЮЧЕНИЕ

В заключении целесообразно обратить внимание на некоторые *фундаментальные* психологические проблемы и риски, тесно связанные с внедрением технологий анализа больших данных в повседневную жизнь общества.

Так, можно ожидать изменений в процессах социального сравнения и социального влияния, так как для личности открывается возможность оценивать себя и других на основе сопоставления с данными о поведении большинства, а сами основания для сравнения становятся более произвольными, легко настраиваемыми с учетом контекста той или иной ситуации. Кроме того, большие данные в “цифровом обществе” превращаются в еще одно основание для формирования стереотипов, когда интерпретация аутгруппового поведения на основе “цифровых следов” оказывается искаженной сформировавшимися ранее предубеждениями или используется для их подтверждения. Большие данные превращаются в критерий оценки личностью социальных ситуаций: как большая или меньшая рельефность этой информации будет сказываться на процессах социального познания? Еще одной важной психологической проблемой становится растущая “прозрачность” жизни человека в деловой и личной сферах. Доступность “цифровых следов” личности другим людям и организациям не только будет служить основанием для формирования новых социальных фобий, но и повлияет на нравственно-психологические компоненты межличностных отношений. Например, можно ожидать изменения критериев доверия и недоверия, формирования новых составляющих социальной зрелости и ответственности личности, когда атрибутивные процессы в межличностном и межгрупповом восприятии основываются не только на уже совершенных поступках, но и на “цифровых следах” намерений. Развитие технологий бесконтактной автоматизированной психологической диагностики существенно повлияет на такие характеристики личности, как готовность к самораскрытию и управление впечатлением, а также формы проявления негативных личностных качеств.

Динамика межличностных отношений, отраженная в “цифровых следах”, станет доступна для более дифференцированной взаимной оценки сторонами, что может повлечь за собой изменения в скорости протекания внутригрупповых процессов. Уже начавшийся переход крупных компаний к непрерывной оценке деятельности своих сотрудников на основе потоковой аналитики создает ряд

психологических проблем в области организационного поведения. Например, встает вопрос о том, как сокращение цикла оценки деятельности сотрудника, а также расширение круга людей и групп, которые могут оценивать ее эффективность, пользуясь потоковой аналитикой, повлияет на мотивацию, приверженность и удовлетворенность работой? Как будут меняться ролевая структура и социально-психологический климат групп в условиях, когда коллектив будет получать свой групповой психологический профиль и обратную связь на свои действия в режиме реального времени, а не с месячной или недельной задержкой? Как командные ментальные модели, включающие в себя представления о работе с большими данными и “цифровых двойниках” объекта управления, будут сказываться на организации совместной деятельности и межгрупповых отношениях?

Большие данные и алгоритмы их обработки “фетишизируются” и уже используются как средство политического воздействия, что приводит к переоценке их реальных возможностей в решении социальных проблем. Фрустрация завышенных ожиданий может вызвать не только кризис доверия к технологии анализа больших данных, но и обрушение доверия к социальным институтам, которые их используют. Формирование рынков данных обострит проблему дефицита межгруппового доверия в отношениях между различными группами горожан, разработчиками систем искусственного интеллекта и технократическими элитами. В эпоху информационных войн с использованием машинного обучения на основе больших данных формирование институционального доверия становится все более трудной социально-психологической задачей. Благодаря большим данным представление общества о самом себе становится более динамичным, появляются новые средства как для верификации социальных представлений, так и для манипулирования ими. Особое значение в этих условиях приобретают цифровая память и социально-психологические механизмы реконструирования коллективного прошлого, отраженного в “цифровых следах” миллионов людей. Можно предположить, что невозможность стереть “цифровые следы” не только индивидуальных, но и групповых поступков повлияет на осмысление коллективного прошлого, настоящего и будущего, а также вызовет к жизни новые формы мнемонических войн.

Дальнейшее развитие интернета вещей и лавинообразное вторжение алгоритмов в повседневную жизнь горожан не только дает в руки психологов новые методы исследований, но и делает невозможным изучение традиционных

социально-психологических феноменов без учета того влияния, которое оказывает на них цифровая среда.

СПИСОК ЛИТЕРАТУРЫ

1. *Агадуллина Е.Р.* Пользователи социальных сетей: современные исследования // Современная зарубежная психология. 2015. Т. 4. № 3. С. 36–46.
2. *Берроуз Р., Севидж М.* После кризиса? Big Data и методологические вызовы эмпирической социологии // Социологические исследования. 2016. № 3 (383). С. 28–35.
3. *Войсунский А.Е.* Психология и искусственный интеллект: новый этап старого взаимодействия // Психология человека как субъекта познания, общения и деятельности / Отв. ред. В.В. Знаков, А.Л. Журавлев. М.: Изд-во “Институт психологии РАН”, 2018. С. 2094–2101.
4. *Журавлев А.Л., Нестик Т.А., Юревич А.В.* Прогноз развития психологической науки и практики к 2030 г. // Психологический журнал. 2016. Т. 37. № 5. С. 45–64.
5. *Журавлев А.Л., Соснин В.А., Китова Д.А., Нестик Т.А., Юревич А.В.* Массовое сознание и поведение: тенденции социально-психологических исследований. М.: Изд-во “Институт психологии РАН”, 2017.
6. *Зуев К.Б., Нестик Т.А.* Основные направления развития психологических исследований: библиометрический анализ по данным Web of Science и статистики поисковых запросов Google // Психологическое знание: Современное состояние и перспективы развития / Под ред. А.Л. Журавлева, А.В. Юревича. М.: Изд-во “Институт психологии РАН”, 2018. С. 671–697.
7. *Корытникова Н.В.* Online Big Data как источник аналитической информации в online-исследованиях // Социологические исследования. 2015. № 8 (376). С. 14–24.
8. *Ледовая Я.А., Тихонов Р.В., Боголюбова О.Н.* Социальные сети как новая среда для междисциплинарных исследований поведения человека // Вестник Санкт-Петербургского университета. Сер. 16. Психология. Педагогика. 2017. Т. 7 (3). С. 193–210.
9. *Ледовая Я.А., Тихонов Р.В., Боголюбова О.Н., Иванов В.Ю., Яминов Б.Р.* Организационно-методические вопросы сбора данных в онлайн-исследованиях поведения пользователей социальной сети “Фейсбук” из России и США // Вестник Санкт-Петербургского университета. Сер. 16. Психология. Педагогика. 2017. Т. 7 (4). С. 308–327. <https://doi.org/10.21638/11701/spbu16.2017.402/> (дата обращения: 12.01.2017).
10. Макропсихология современного российского общества / Под ред. А.Л. Журавлева, А.В. Юревича. М.: Изд-во “Институт психологии РАН”, 2009.
11. *Михеев Е.А., Нестик Т.А.* Дезинформация в социальных сетях: состояние и перспективы психологических исследований // Фундаментальные и прикладные исследования современной психологии: результаты и перспективы развития / Отв. ред. А.Л. Журавлев, В.А. Кольцова. М.: Изд-во “Институт психологии РАН”, 2017. С. 2021–2030.
12. *Нестик Т.А.* Развитие цифровых технологий и будущее психологии // Вестник Московского государственного областного университета. Сер. Психологические науки. 2017. № 3. С. 6–15.
13. *Погорский Э.К.* Особенности цифровых гуманитарных наук // Знание. Понимание. Умение. 2014. № 5. URL: http://www.zpu-journal.ru/e-zpu/2014/5/Pogorskiy_Digital-Humanities/ (дата обращения: 12.01.2017).
14. *Смирнов И.В.* Введение в анализ естественных языков: учебно-методическое пособие. М.: Изд-во РУДН, 2014.
15. *Солдатова Г.У., Рассказова Е.И., Нестик Т.А.* Цифровое поколение России: компетентность и безопасность. М.: Смысл, 2017.
16. *Тищенко В.И.* Информатика сообществ: Противоречия становления // Социальные сети и виртуальные сетевые сообщества / Отв. ред. Л.Н. Верченнов, Д.В. Ефременко, В.И. Тищенко. М., 2013. С. 11–18.
17. *Толстова Ю.Н.* Социология и компьютерные технологии // Социологические исследования. 2015. № 8 (376). С. 3–13.
18. *Al-Mosaiwi M., Johnstone T.* In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation // Clinical Psychological Science January 5, 2018 URL: <https://doi.org/10.1177/2167702617747074> (дата обращения: 01.03.2017).
19. *Azucar D., Marengo D., Settanni M.* Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis // Personality and Individual Differences. 2018. V. 124. P. 150–159.
20. *Bessi A.* Personality traits and echo chambers on Facebook // Computers in Human Behavior. 2016. V. 65. P. 319–324.
21. *Burnap P., Williams M.L., Sloan L.* Tweeting the terror: Modelling the social media reaction to the Woolwich terrorist attack // Social Network Analysis and Mining. 2014. V. 4 (1). P. 1–14.
22. *Candia C., Jara-Figueroa C., Rodriguez-Sickert C., Barabási A.-L., Hidalgo C.A.* The universal decay of collective memory and attention // Nature Human Behaviour. 2019. V. 3. P. 82–91.

23. *Crump M.J.C., McDonnell J.V., Gureckis T.M.* Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research // PLoS ONE. 2013. 8(3): e57410.
24. *Deign J.* Big data dating: romance goes digital // The Network. Cisco's Technology News Site. September 13, 2017. URL: <https://newsroom.cisco.com/feature-content?articleId=1875523> (дата обращения: 01.03.2017).
25. *Guzzo R.A., Fink A.A., King E., Tonidandel S., Landis R.S.* Big data recommendations for industrial-organizational psychology // Industrial and Organizational Psychology. 2015. V. 8 (4). P. 491–508.
26. *Jones M.N.* Developing cognitive theory by mining large-scale naturalistic data // Big Data in Cognitive Science. Frontiers of Cognitive Psychology / Ed. M.N. Jones. N.Y.: Routledge, 2017. P. 1–12.
27. *Karlsson Linner R., Biroli P., Kong E., [...] Beauchamp J.P.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences // Nature Genetics. 2019. V. 51 (2). P. 245–257.
28. *Kosinski M.* Computational Psychology // Advanced social psychology: the state of the science. 2nd Ed. / Eds. E.J. Finkel and R.F. Baumeister. Oxford: Oxford University Press, 2019. P. 499–523.
29. *Kosinski M., Stillwell D., Graepel T.* Private traits and attributes are predictable from digital records of human behavior // Proceedings of the National Academy of Sciences (PNAS). 2013. V. 110. № 15. P. 5802–5805.
30. *Kosinski M., Wang Y., Lakkaraju H., Leskovec J.* Mining Big Data to Extract Patterns and Predict Real-Life Outcomes // Big Data in Psychology. Special issue of Psychological Methods. 2016. V. 21. № 4. P. 493–506.
31. *Kramer A.D., Guillory J.E., Hancock J.T.* Experimental evidence of massive-scale emotional contagion through social networks // Proceedings of the National Academy of Sciences. 2014. V. 111 (24). P. 8780–8790.
32. *Lambiotte R., Kosinski M.* Tracking the Digital Footprints of Personality // Proceedings of the Institute of Electrical and Electronics Engineers (IEEE). 2014. V. 102. № 12. P. 1934–1939.
33. *Manovich L.* Automating Aesthetics: Artificial Intelligence and Image Culture. 2017. URL: http://manovich.net/content/04-projects/101-automating-aesthetics-artificial-intelligence-and-image-culture/automating_aesthetics.pdf (дата обращения: 01.03.2017).
34. *Mitroff S.R., Biggs A.T., Adamo S.H., Dowd E.W., Winkle J., Clark K.* What can 1 billion trials tell us about visual search? // Journal of Experimental Psychology: Human Perception and Performance. 2015. V. 41. P. 1–5.
35. *Moliner P.* Psychologie sociale de l'image. Grenoble: Presses Universitaires de Grenoble, 2016.
36. *Monteith S., Glenn T., Geddes J., Bauer M.* Big data are coming to psychiatry: a general introduction // International Journal of Bipolar Disorders. 2015. V. 3. P. 21.
37. *Obschonka M., Stuetzer M., Rentfrow P.J., Shaw-Taylor L., Satchell M., Silbereisen R.K., Potter J., Gosling S.D.* In the Shadow of Coal: How Large-Scale Industries Contributed to Present-Day Regional Differences in Personality and Well-Being // Journal of Personality and Social Psychology. 2018. V. 115 (5). P. 903–927.
38. *Pettit M.* Historical time in the age of Big Data: Cultural psychology, historical change, and the Google Books Ngram Viewer // History of Psychology. 2016. V. 19 (2). P. 141–153.
39. *Putka D.J., Oswald F.L.* Implications of the Big Data movement for the advancement of IO science and practice // Big data at work: The data science revolution and organizational psychology / Eds. Tonidandel S., King E., Cortina J.M.. N.Y., NY: Routledge, 2015. P. 181–212.
40. *Rudder C.* Dataclysm: Who We Are (When We Think No One's Looking). N.Y.: Crown, 2014.
41. *Salganik M.J.* Bit by Bit: Social Research in the Digital Age. Princeton & Oxford: Princeton University Press, 2018.
42. *Spitzberg B.H.* Toward a model of meme diffusion (M 3 D) // Communication Theory. 2014. V. 24 (3). P. 311–339.
43. *Stewart R., Davis K.* 'Big Data' in mental health research: current status and emerging possibilities // Social Psychiatry and Psychiatric Epidemiology. 2016. V. 51. № 8. P. 1055–1072.
44. *Vosoughi S., Roy D., Aral S.* The spread of true and false news online // Science. 2018. V. 359. № 6380. P. 1146–1151.

BIG DATA ANALYSIS IN PSYCHOLOGY AND SOCIAL SCIENCES: PERSPECTIVE DIRECTIONS OF RESEARCH²

T. A. Nestik*, A. L. Zhuravlev**

*Institute of Psychology, Russian Academy of Sciences;
129366, Moscow, Yaroslavskaya str., 13, Russia.*

**ScD (psychology), professor of RAS, acting head of the laboratory of social and economic psychology.
E-mail: nestik@gmail.com*

***Member of RAS, scientific leader of the Federal State-Financed Institution,
Institute of Psychology RAS; ScD (psychology), professor.
E-mail: alzhuravlev2018@yandex.ru*

Received 29.08.2019

Abstract. The consequences of Big Data revolution for psychology and social sciences are analyzed. The article presents an overview of the four types of psychological research based on the Big Data analysis: secondary data analysis; crowdsourcing projects and matching of personal digital footprints of social network users with data of psychometric questionnaires; research of psychological phenomena based on network analysis, natural language processing and machine learning without surveys; natural experiments and data gathering by portable gadgets. A special attention is paid to psychological profiling of personality by digital footprints. The perspective directions for further research of dynamics in socio-psychological phenomena at the personal, interpersonal, group, and macro-social levels with applying computer linguistic and visual data analysis are proposed. The psychological and social risks of Big Data based technologies adoption in the everyday life of society are discussed.

Keywords: Big Data, digital footprints, visual data, social networks, group dynamics, macro-psychological indicators, risks.

REFERENCES

1. Agadullina E.R. Pol'zovateli social'nyh setej: sovremennye issledovaniya // *Sovremennaja zarubezhnaja psihologija*. 2015. V. 4. № 3. P. 36–46. (in Russian)
2. Berrouz R., Sevidzh M. Posle krizisa? Big Data I metodologicheskie vyzovy empiricheskoy sociologii // *Sociologicheskie issledovaniya*. 2016. № 3 (383). P. 28–35. (in Russian)
3. Vojskunsij A.E. Psihologija i iskusstvennyj intellekt: novyj etap starogo vzaimodejstviya // *Psihologija cheloveka kak sub'ekta poznaniya, obshheniya i dejatel'nosti* / Eds. V.V. Znakov, A.L. Zhuravljov. Moscow: Izd-vo "Institut psihologii RAN", 2018. P. 2094–2101. (in Russian)
4. Zhuravlev A.L., Nestik T.A., Jurevich A.V. Prognoz razvitiya psihologicheskoy nauki i praktiki k 2030 g. // *Psihologicheskij zhurnal*. 2016. V. 37. № 5. P. 45–64. (in Russian)
5. Zhuravlev A.L., Sosnin V.A., Kitova D.A., Nestik T.A., Jurevich A.V. Massovoe soznaniye i povedeniye: tendencii social'no-psihologicheskikh issledovanij. Moscow: Izd-vo "Institut psihologii RAN", 2017. (in Russian)
6. Zuev K.B., Nestik T.A. Osnovnye napravleniya razvitiya psihologicheskikh issledovanij: bibliometricheskij analiz po dannym Web of Science I statistiki poiskovykh zaprosov Google // *Psihologicheskoe znanie: Sovremennoe sostojanie i perspektivy razvitiya* / Eds. A.L. Zhuravlev, A.V. Jurevich. Moscow: Izd-vo "Institut psihologii RAN", 2018. P. 671–697. (in Russian)
7. Korytnikova N.V. Online Big Data kak istochnik analiticheskoy informacii v online-issledovanijah // *Sociologicheskie issledovaniya*. 2015. № 8 (376). P. 14–24. (in Russian)
8. Ledovaja Ja.A., Tihonov R.V., Bogoljubova O.N. Social'nye seti kak novaja sreda dlja mezhdisciplinarnykh issledovanij povedeniya cheloveka // *Vestnik Sankt-Peterburgskogo universiteta. Seriya 16. Psihologija. Pedagogika*. 2017. V. 7 (3). P. 193–210. (in Russian)
9. Ledovaja Ja.A., Tihonov R.V., Bogoljubova O.N., Ivanov V. Ju., Jaminov B.R. Organizacionno-metodicheskie voprosy sbora dannyh v onlajn-issledovanii povedeniya pol'zovatelej social'noj seti "Fejsbuk" iz Rossii i SShA // *Vestnik Sankt-Peterburgskogo universiteta. Seriya 16. Psihologija. Pedagogika*. 2017. V. 7 (4). P. 308–327. <https://doi.org/10.21638/11701/spbu16.2017.402/> (data obrashheniya: 12.01.2017). (in Russian)
10. Makropsihologija sovremennogo rossijskogo obshhestva / Eds. A.L. Zhuravlev, A.V. Jurevich. Moscow: Izd-vo "Institut psihologii RAN", 2009. (in Russian)

² The research was supported by RSF grant № 18-18-00439.

11. *Mihev E.A., Nestik T.A.* Dezinformacija v social'nyh set-jah: sostojanie i perspektivy psihologicheskikh issledovanij // Fundamental'nye i prikladnye issledovanija sovremennoj psihologii: rezul'taty i perspektivy razvitija / Eds. A.L. Zhuravljov, V.A. Kol'cova. Moscow: Izd-vo "Institut psihologii RAN", 2017. P. 2021–2030. (in Russian)
12. *Nestik T.A.* Razvitie cifrovyyh tehnologij i budushhee psihologii // Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Psihologicheskie nauki. 2017. № 3. P. 6–15. (in Russian)
13. *Pogorskiy Je.K.* Osobennosti cifrovyyh gumanitarnyyh nauk // Znanie. Ponimanie. Umenie. 2014. № 5. URL: http://www.zpu-journal.ru/e-zpu/2014/5/Pogorskiy_Digital-Humanities/ (data obrashheniya: 12.01.2017). (in Russian)
14. *Smirnov I.V.* Vvedenie v analiz estestvennykh jazykov: uchebno-metodicheskoe posobie. Moscow: Izd-vo RUDN, 2014. (in Russian)
15. *Soldatova G.U., Rasskazova E.I., Nestik T.A.* Cifrovoe pokolenie Rossii: kompetentnost' i bezopasnost'. Moscow: Smysl, 2017. (in Russian)
16. *Tishhenko V.I.* Informatika soobshhestv: Protivorechija stanovleniya // Social'nye seti i virtual'nye setevye soobshhestva / Eds. L.N. Verchenov, D.V. Efremenko, V.I. Tishhenko. Moscow, 2013. P. 11–18. (in Russian)
17. *Tolstova Ju.N.* Sociologiya i komp'yuternye tehnologii // Sociologicheskie issledovanija. 2015. № 8 (376). P. 3–13. (in Russian)
18. *Al-Mosaiwi M., Johnstone T.* In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation // Clinical Psychological Science January 5, 2018. URL: <https://doi.org/10.1177/2167702617747074> (data obrashheniya: 01.03.2017).
19. *Azucar D., Marengo D., Settanni M.* Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis // Personality and Individual Differences. 2018. V. 124. P. 150–159.
20. *Bessi A.* Personality traits and echo chambers on Facebook // Computers in Human Behavior. 2016. V. 65. P. 319–324.
21. *Burnap P., Williams M.L., Sloan L.* Tweeting the terror: Modelling the social media reaction to the Woolwich terrorist attack // Social Network Analysis and Mining. 2014. V. 4 (1). P. 1–14.
22. *Candia C., Jara-Figueroa C., Rodriguez-Sickert C., Barabási A.-L., Hidalgo C.A.* The universal decay of collective memory and attention // Nature Human Behaviour. 2019. V. 3. P. 82–91.
23. *Crump M.J.C., McDonnell J.V., Gureckis T.M.* Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research // PLoS ONE. 2013. 8 (3): e57410.
24. *Deign J.* Big data dating: romance goes digital // The Network. Cisco's Technology News Site. September 13, 2017. URL: <https://newsroom.cisco.com/feature-content?articleId=1875523> (data obrashheniya: 01.03.2017).
25. *Guzzo R.A., Fink A.A., King E., Tonidandel S., Landis R.S.* Big data recommendations for industrial-organizational psychology // Industrial and Organizational Psychology. 2015. V. 8 (4). P. 491–508.
26. *Jones M.N.* Developing cognitive theory by mining large-scale naturalistic data // Big Data in Cognitive Science. Frontiers of Cognitive Psychology / Ed. M.N. Jones. N.Y.: Routledge, 2017. P. 1–12.
27. *Karlsson Linner R., Biroli P., Kong E., [...] Beauchamp J.P.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences // Nature Genetics. 2019. V. 51 (2). P. 245–257.
28. *Kosinski M.* Computational Psychology // Advanced social psychology: the state of the science. 2nd Ed. / Eds. E.J. Finkel and R.F. Baumeister. Oxford: Oxford University Press, 2019. P. 499–523.
29. *Kosinski M., Stillwell D., Graepel T.* Private traits and attributes are predictable from digital records of human behavior // Proceedings of the National Academy of Sciences (PNAS). 2013. V. 110. № 15. P. 5802–5805.
30. *Kosinski M., Wang Y., Lakkaraju H., Leskovec J.* Mining Big Data to Extract Patterns and Predict Real-Life Outcomes // Big Data in Psychology. Special issue of Psychological Methods. 2016. V. 21. № 4. P. 493–506.
31. *Kramer A.D., Guillory J.E., Hancock J.T.* Experimental evidence of massive-scale emotional contagion through social networks // Proceedings of the National Academy of Sciences. 2014. V. 111 (24). P. 8780–8790.
32. *Lambiotte R., Kosinski M.* Tracking the Digital Footprints of Personality // Proceedings of the Institute of Electrical and Electronics Engineers (IEEE). 2014. V. 102. № 12. P. 1934–1939.
33. *Manovich L.* Automating Aesthetics: Artificial Intelligence and Image Culture. 2017. URL: http://manovich.net/content/04-projects/101-automating-aesthetics-artificial-intelligence-and-image-culture/automating_aesthetics.pdf (data obrashheniya: 01.03.2017).
34. *Mitroff S.R., Biggs A.T., Adamo S.H., Dowd E.W., Winkle J., Clark K.* What can 1 billion trials tell us about visual search? // Journal of Experimental Psychology: Human Perception and Performance. 2015. V. 41. P. 1–5.
35. *Moliner P.* Psychologie sociale de l'image. Grenoble: Presses Universitaires de Grenoble, 2016.
36. *Monteith S., Glenn T., Geddes J., Bauer M.* Big data are coming to psychiatry: a general introduction // International Journal of Bipolar Disorders. 2015. V. 3. P. 21.
37. *Obschonka M., Stuetzer M., Rentfrow P.J., Shaw-Taylor L., Satchell M., Silbereisen R.K., Potter J., Gosling S.D.* In the Shadow of Coal: How Large-Scale Industries Contributed to Present-Day Regional Differences in Personality and Well-Being // Journal of

- Personality and Social Psychology. 2018. V. 115 (5). P. 903–927.
38. *Pettit M.* Historical time in the age of Big Data: Cultural psychology, historical change, and the Google Books Ngram Viewer // *History of Psychology*. 2016. V. 19 (2). P. 141–153.
39. *Putka D.J., Oswald F.L.* Implications of the Big Data movement for the advancement of IO science and practice // *Big data at work: The data science revolution and organizational psychology* / Eds. Tonidandel S., King E., Cortina J.M.. New York, NY: Routledge, 2015. P. 181–212.
40. *Rudder C.* Dataclysm: Who We Are (When We Think No One's Looking). N.Y.: Crown, 2014.
41. *Salganik M.J.* Bit by Bit: Social Research in the Digital Age. Princeton & Oxford: Princeton University Press, 2018.
42. *Spitzberg B.H.* Toward a model of meme diffusion (M 3 D) // *Communication Theory*. 2014. V. 24 (3). P. 311–339.
43. *Stewart R., Davis K.* 'Big Data' in mental health research: current status and emerging possibilities // *Social Psychiatry and Psychiatric Epidemiology*. 2016. V. 51. № 8. P. 1055–1072.
44. *Vosoughi S., Roy D., Aral S.* The spread of true and false news online // *Science*. 2018. V. 359. № 6380. P. 1146–1151.