

Научная статья

УДК 316.1

doi: 10.22394/2079-1690-2022-1-2-247-251

BIG DATA: НОВЫЕ ОРГАНИЗАЦИОННЫЕ ВОЗМОЖНОСТИ И СОЦИАЛЬНЫЕ РИСКИ

Лейла Шамильевна Крупеникова¹, Владимир Иванович Курбатов²

¹Южный федеральный университет, Ростов-на-Дону, Россия, leilakrupenikova@gmail.com

²Южно-Российский филиал Федерального научно-исследовательского социологического центра Российской академии наук, Ростов-на-Дону, Россия, kurbahsy@list.ru

Аннотация. Статья посвящена аналитическому обзору результатов применения «Big Data» («Большие Данные») различными международными компаниями и обзору экспертных оценок, в соответствии с чем выявлены возможности использования больших массивов данных информации и, одновременно с этим показано, что возникли новые социокультурные вызовы, проблемы, связанные с информацией, которая может подбираться по смещенным выборкам, социальные риски, обусловленные использованием приватной информации и личных данных не по назначению, в частности для создания ложных новостей, искаженной официальной статистики, для фальсификации опросов и выборов.

Ключевые слова: Big Data, «Большие данные», цифровизация, безопасность информации и бизнеса, организационные возможности и социальные риски использования больших массивов информации, социокультурные проблемы

Для цитирования: Крупеникова Л. Ш., Курбатов В. И. Big Data: новые организационные возможности и социальные риски // Государственное и муниципальное управление. Ученые записки. 2022. № 2. С. 247–251. <https://doi.org/10.22394/2079-1690-2022-1-2-247-251>

Sociology Problems

Original article

BIG DATA: NEW ORGANIZATIONAL OPPORTUNITIES AND SOCIAL RISKS

Leila Sh. Krupenikova, Vladimir I. Kurbatov

¹Southern Federal University, Rostov-on-Don, Russia, leilakrupenikova@gmail.com

²South-Russian Branch of Federal Research Sociological Center of the Russian Academy of Sciences, Rostov-on-Don, Russia, kurbahsy@list.ru

Abstract. The article is devoted to an analytical review of the results of the application of "Big Data" ("Big Data") by various international companies and a review of expert assessments, in accordance with which the possibilities of using large data arrays as sociological information and, at the same time, sociocultural challenges, problems associated with information, which can be selected from biased samples, the risks associated with the use of private information and personal data for other purposes, in particular to create false news, distorted official statistics, to falsify polls and elections.

Keywords: Big Data, big data sets, models of social reality, digitalization, digital sociology

For citation: Krupenikova L. Sh., Kurbatov V. I. Big Data: new organizational opportunities and social risks. *State and Municipal Management. Scholar Notes. 2022;(2):247–251. (In Russ.).* <https://doi.org/10.22394/2079-1690-2022-1-2-247-251>.

Big Data «Большие данные» – это термин, предназначенный для указания на систематически структурированные и бессистемные неструктурированные разнообразные данные, которые характеризуются огромным и постоянно пополняемым объемом различной многопрофильной информации, которая обрабатывается с помощью специальных программных инструментов и технологий.

Термин «Big Data» возник и стал широко использоваться сравнительно недавно, во всяком случае исследователи связывают появление данного термина с именем эксперта Клиффорда Линча, который в журнале «Nature» в своей статье так назвал накопление информации. С тех пор этот термин стал общепринятым и по оценкам Американской ассоциации изучения общественного мнения, характеризующей статьи таких авторов, как М. Берг, П. Биенер, П. Декер, Л. Джэпес, Ф. Креутер, К. Лэмп, Дж. Лэйн, О. Нейл и других, определяет данное понятие как описание огромных по объему и самых разнообразных по содержанию, наполнению и характеристикам компьютерных практик, технологических и технических приемов, различной информации, касающейся производства и бизнеса, народонаселения и политики, образования и здравоохранения, науки и искусства, этики и права, миграции и демографии, равно и проблем, которые связаны с разными сторонами общественной жизни [1–2].

Информационный объем «Больших данных» постоянно увеличивается и по данным аналитиков, если в 2003 г. он составлял пять эксабайтов данных (1 ЭБ составляет 1 миллиард гигабайтов), в 2020 г. – сорок четыре зеттабайта (1 ЗБ составляет 1024 эксабайта), то в 2025 г. по прогнозам этот объем вырастет на порядок. Источниками «Больших данных» являются блоги, социальные сети, сайты, форумы, СМИ, корпоративные архивы, многообразные транзакции, различные профессиональные базы данных, в том числе показания разнообразных считывающих устройств, данные геолокации, метеорологии, сотовой связи т.д.

По экспертным оценкам исследовательской программы «Cisco Annual Internet Report» принцип работы «Больших данных» может быть объяснен тремя «V» («volume», «velocity» и «variety»), соответственно обозначающими физический объем, скорость получения новых данных и их быстрая обработка¹. В дальнейшем, как пишут С. Барокас, С. Бендер, Р. Китчин, Х. Ниссенбаум и В. Стодден, принцип трех «V» расширился до восьми «V», включив в себя «veracity» (достоверность), «viability» (жизнеспособность), «value» (ценность), «variability» (переменчивость) и «visualization» (визуализация) [3–4]. Появились и получили динамичное развитие конкретизированные «Большие данные», например, «Data Science», «Data Busines», «Data Mining» и другие. Международная компания «McKinsey» характеризуют методы обработки данных следующим образом: краудсорсинг (классификация), интеграция данных (data fusion and integration), машинное обучение, сетевой анализ с использованием генетических алгоритмов (genetic algorithm), распознавание образов, прогнозная аналитика, имитационное моделирование (simulation), пространственный анализ (spatial analysis), статистический анализ, визуализация аналитических данных.

По данным аналитиков, таких как С. Белфки, Ф.-З. Бенджелоун, А. А. Лахцен и А. Оссоуз 59,4 % респондентов заявили, что использование «Больших данных» повышают производительность работы, а 66,7% заявили, что это помогло сократить расходы. Подавляющее большинство (92 %) представителей компаний одобряют возможности использования «Больших данных», а 89 % полагают, что это дало дополнительный ресурс их бизнесу для совершенствования и развития [5]. В целом, оценивая использование «Больших данных» дает возможность принятия лучших решений, помогает внедрять инновации, повышает производительность, помогает находить клиентов и рынки и создает условия обеспечения безопасности корпоративной информации. Масштабность внедрения «Больших данных» дали основания Дж. Ван Дейку даже ввести в оборот новые термины – «датафикация» и датаизм» [6].

По данным опроса «Cisco Connected World Technology Report» источниками оперативных данных являются для пользователей являются: 74 % собирают соответствующие данные; 55 % собирают данные по истории; 48 % снимают данные с мониторов и датчиков; 40% используют

¹ Cisco Annual Internet Report. White Paper (2018–2023).

<http://www.cisco.com/en/US/netsol/ns1204/index.html#~Products> (Date of treatment – 03/10/2018); Economist Intelligence Unit survey data // <https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html> (accessed 9.03.2020).

данные, соответствующие реальному времени¹. Исследование «Accenture», в котором участвовали более тысячи руководителей компаний из 19 стран мира реализовалось в отчете «Intelligence Unit», который показал, что основные преимущества «Больших данных» 56 % усматривают в возможности поиска и получения новых источников дохода, 51 % отмечают приобретение нового опыта в работе с информацией о клиентах, 50 % видят преимущество в реализации возможности производства новых товаров и услуг и 47 % отмечают, что благодаря использованию больших массивов информации удалось обеспечить приток новых клиентов.

Для большинства, опрошенных, которых оказалось 51 %, очень важной проблемой является проблема безопасности информации и бизнеса; для 47 % это проблемы бюджета, для 41 % нехватка квалифицированных кадров, способных работать с большими объемами информации и для 35 % проблемой оказывается интеграция информации в сложившуюся бизнес-систему. Для 73 % респондентов «Большие данные» это «локомотив прогресса», при этом они несколько расходятся в оценках того, что считать «Большими данными»: 65 % полагают, что это «соответствующие каталоги данных», для 60 % это прежде всего «продвинутая аналитика», а для 50 % респондентов это «данные инструментов визуализации»².

Результаты исследования международной компании «New Vantage Partners» в области «Больших данных» для руководителей, проведенное в 2018 году, проанализированное экспертами, показали, что 97,2 % опрошенных фирм инвестируют в сферу «Больших данных». По данным отчета об опросе, проведенного «New Vantage Partners» 36,2 % считают, что это лучшие инвестиции, при этом уже 84,1 % начали работать над проектами, связанными с «Большими данными», а 59,0 % уже добились определенного результата, который делает инвестиции оправданными. 69,0 % респондентов утверждают, что благодаря таким инвестициям их компании успешно конкурируют и растут, а более половины респондентов (54,7 %) утверждают, что использование инструментов «Больших данных» дает увеличение дохода и ускорение роста. Более того, согласно отчету 41,7 % респондентов заявили, что применение технологий обработки больших массивов информации повышает гибкость их бизнеса [7–8].

Таким образом, по оценкам экспертов «Большие данные» дают возможность принимать лучшие решения, а инструменты работы с массивом большой информации дают возможность просматривать больше информации и на более высоких скоростях, что помогает лучше структурировать входящую информацию. Что касается организации бизнеса, то «Большие данные» помогают сократить расходы организаций, увеличить доход, улучшить обслуживание клиентов. Это связано с тем, что доступ к аналитике операций по сбору данных упрощает обслуживание клиентов. Это помогает завоевать доверие потенциальных клиентов. Компании также могут быстрее выходить на рынок с новыми продуктами или услугами. «Большие данные» создают условия для внедрения инноваций, чем обеспечивается продвинутость компании и гибкость ее поведения на рынке, что способствует развитию электронной коммерции.

«Большие данные» привнесли в общественную жизнь не только новые ресурсы и новые возможности, но также и новые вызовы, проблемы и риски. Представляется возможным, основываясь на экспертных оценках, выделить несколько основных групп вызовов, проблем и рисков. Первая группа – это проблемы технологические и организационные. Так, по данным исследования «New Vantage Partners», ориентированного на анализ использования «Больших данных» для трансформации бизнеса, было отмечено, что 90,9 % опрошенных компаний указывают на то, что самой большими проблемами являются трудности управления большими объемами информации, что добавляют организационные проблемы, отсутствие кадров и соответствующего опыта работников [7–8].

¹ Cisco Annual Internet Report. White Paper (2018–2023).

<http://www.cisco.com/en/US/netsol/ns1204/index.html#~Products> (Date of treatment – 03/10/2018); Economist Intelligence Unit survey data // <https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html> (accessed 9.03.2020)

² Big Data Analytics Market. Companies profiled In Big Data Analytics Market Microsoft Corporation (Washington, U.S.), Dun & Bradstreet Holdings, Inc. (New Jersey, U.S.), Equifax, Inc. (Georgia, U.S.), IBM Corporation (New York, U.S.), Salesforce.com, Inc. (California, U.S.), SAP SE (Walldorf, Germany), Verisk Analytics Inc. (New Jersey, U.S.), Dow Jones & Company, Inc. (New York, U.S.), Moody's Corporation (New York, U.S.), Oracle Corporation (Texas, U.S.), SAS Institute Inc. (North Carolina, U.S.), QlikTech International AB (Pennsylvania, U.S.), TransUnion (Illinois, U.S.), Fair Isaac Corporation (California, U.S.), Experian Information Solutions, Inc. (California, U.S.).

Другая группа проблем связана с безопасностью, а именно с тем, как используется информация «Больших данных», не приведет ли она к вторжению в частную область и некорректному использованию личных данных, на что указывают многие эксперты, в частности, В. Вейхонг, Г. фон Войгт, Х. Кай, С. Сзонгот, Б. Хенне и другие. Во всяком случае, эксперты указывают, что анонимизация пользователей стала весьма проблематичной, что порождает целый ряд этических проблем, связанных с вторжением государства и корпораций в частную область. Эксперты отмечают, что аналитика больших массивов информации сделалась привлекательной мишенью для кибератак, в частности, согласно отчету Risk-Based Security Mid-Year «Большие данные», уже в первой половине 2019 г. в результате утечек данных было раскрыто 4,1 млрд записей, от чего пострадала конфиденциальность и был нанесен вред клиентам и партнерам [9–10].

В других случаях, государство, организации или группы лиц могут сами использовать частную информацию не по назначению. Это может быть и то, как отмечают Р. Джозеф и Н. Джонсон и другие эксперты, что называется «цифровым концлагерем», «цифровым капитализмом» или просто государственной слежкой, может быть вмешательством в выборы, распространением фальшивых новостей, манипулированием, использованием для создания ложного имиджа, например, препятствующего карьерному росту и продвижению по службе. Эксперты прямо называют, что аналитика «Больших данных» может способствовать дискриминации по самым разным признакам. Все это вышесказанное воспроизводит цифровое неравенство [11–12].

Среди разных вызовов и социальных рисков использования «Больших данных» эксперты называют социокультурные проблемы, характеризующие необходимость адаптации к новой цифровой среде, например, в опросе «New Vantage» о достоинствах и недостатках «Больших данных» среди руководителей фирм и компаний, почти 99 % руководителей заявили, что их компании находятся только в процессе создания новой культуры и, как отмечают Д. Бойд и К. Крауфорд, только около трети персонала адаптировались в новой среде. При этом, во многом достоинства и возможности «Больших данных» зависят от качества этих данных, т.е. от достоверности и оперативности информации, о чем эксперты пишут, что она не всегда точна, порою тенденциозно подобрана, либо ее выборка сделана по произвольным критериям, а та информация, которая контролируется правительствами, например, официальная статистика подается таким образом, каким это выгодно заказчику [13].

В результате аналитического обзора оценок применения «Больших данных» различными международными компаниями, наряду с положительными оценками, о которых было сказано выше, выделены несколько основных групп вызовов, проблем и социальных рисков использования «Больших данных»:

- технологические и организационные проблемы, связанные с трудностями управления большими массивами информации;
- недостаточная подготовленность кадров к работе с операционными системами;
- этические проблемы, связанные с использованием частной информации и личных данных;
- проблемы социальных рисков тотального контроля за цифровой личностью со стороны государства, корпораций и групп людей, которые могут вторгаться в частную область и использовать персональные данные не по назначению;
- социокультурные проблемы, характеризующие необходимость адаптации людей к новой цифровой среде и трудности формирования новой цифровой культуры.
- проблемы, связанные с недостоверной, непроверенной, неточной или неполной информацией, которая может подбираться по смещенным выборкам, тенденциозно и субъективно, не быть объективной и оперативной;
- проблемы применения «Больших данных» не по назначению, в частности, для ложных новостей, искаженной официальной статистики, для фальсификации опросов и выборов.

Список источников / References

1. Lynch, C.A. Big data: How do your data grow? *Nature*. 2008; 455(7209):28-41. DOI:10.1038/455028a.
2. Japac, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., O'Neil, C., Usher, A. Big Data in Survey Research: AAPOR Task Force Report Get access Arrow. *Public Opinion Quarterly*. 2015;79(4):839–880. Available from: <https://doi.org/10.1093/poq/nfv039> [Accessed 3.11.2015].

3. Barocas, S., Nissenbaum, H. «Big Data's End Run around Anonymity and Consent». 2014;44-75. *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, edited by J. Lane, V. Stodden, S. Bender, and H. Nissenbaum. Cambridge: Cambridge University Press.
4. Kitchin R. Big Data and Human Geography: Opportunities, Challenges and Risks. *Dialogues in Human Geography*. 2013;3(3):262–267.
5. Oussous, A., Benjelloun, F.-Z., Lahcen, A.A., Belfki, S. Projects: Big Data Big Data analysis and security. *Journal of King Saud University*. 2017;30(4):1-18. DOI:10.1016/j.jksuci.2017.06.001
6. Van Dijck J. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*. 2014;12 (2):197–208.
7. Davenport T.H., Bea R. How Big Data and AI are Accelerating Business. *Data and Innovation*. 2019. Available from: www.newvantage.com.
8. Akter S., Michael K., Uddin M., McCarth, G. Transforming business using digital innovations: the application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*. 2022;(308):59.
9. Hongbing, C., Chunming, R., Kai, H., Weihong, W., Yanyan, L. Secure big data storage and sharing scheme for cloud tenants. *Communications, China*. 2015;12(6):106–115. doi: 10.1109/CC.2015.712246.
10. Szongott C., Henne B., von Voigt G. Big data privacy issues in public social media. In: 6th IEEE international conference on digital ecosystems technologies. 2012:1-6.
11. Joseph R. C., Johnson N. A. Big data and transformational government. *IT Professional*. 2013;15 (6):43–48.
12. Krishnamurthy R., Desouza K.C. Big data analytics: the case of the social security administration. *Information Polity*. 2014;19(3/4):165–178.
13. Boyd D., Crawford K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*. 2012;15 (5):662–679.

Информация об авторах

Л. Ш. Крупеникова – канд. социол. наук, доц. кафедры прикладной социологии Института социологии и регионоведения ЮФУ;

В. И. Курбатов – докт. филос. наук, проф., гл. научн. сотр. Южно-Российского филиала Федерального научно-исследовательского социологического центра Российской академии наук.

Information about the authors

L. Sh. Krupenikova – Candidate of Sociological Sciences, Associate Professor of Department of Applied Sociology of Southern Federal University.

V. I. Kurbatov – Doctor of Philosophical Sciences, Professor, Chief Researcher of South-Russian branch of Federal Research Sociological Center of the Russian Academy of Sciences.

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации.

Авторы заявляют об отсутствии конфликта интересов.

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Статья поступила в редакцию 13.05.2022; одобрена после рецензирования 27.05.2022; принята к публикации 30.05.2022.

The article was submitted 13.05.2022; approved after reviewing 27.05.2022; accepted for publication 30.05.2022.