



DOI: 10.22363/2312-8313-2022-9-2-150-161

Научная статья / Research article

Цифровая трансформация и большие данные

С.С. Кузора¹  , И.П. Натаров² 

¹Дальневосточный федеральный университет,
690091, Россия, Владивосток, ул. Суханова, 8

²Министерство экономического развития Приморского края,
690110, Россия, Владивосток, ул. Светланская, 22

 webkuzora@mail.ru

Аннотация. На сегодняшний день Российская Федерация уделяет особое внимание цифровой трансформации страны в целом и проявляет отдельный интерес к большим данным в частности, что подтверждается рядом проанализированных документов и сложившейся конъюнктурой в предметной области. Исследование посвящено анализу больших данных, одной из областей использования которых является государство, а точнее работа федеральных и региональных органов исполнительной власти. Основная ценность больших данных с позиции государства видится в возможности работы над объемами разнородной информации с целью повышения эффективности в принятии управленческих решений по широкому кругу вопросов. Основной целью исследования является распространение опыта анализа больших данных, которые используются в работе Министерства экономического развития Приморского края. Для этого предлагается обратиться к теоретическим аспектам больших данных, узнать основную цепочку работ над ними, а также обратить внимание на отечественный и зарубежный опыт использования данных в некоторых областях знаний. Практической частью исследования является описание опыта работы Минэкономразвития ПК, имеющее непосредственное отношение к цифровому развитию региона и осуществляющее определенный перечень работ с разнородными данными. В качестве примера обращается внимание на апробированный подход по анализу большого набора открытых данных, характеризующих работу и развитие малого и среднего предпринимательства в Приморском крае, размещенных на сайте Федеральной налоговой службы Российской Федерации. В заключении подчеркивается, что предлагаемый в исследовании подход по работе с данными может быть адаптирован под схожие потребности в других федеральных и региональных органах власти.

Ключевые слова: большие данные, анализ больших данных, цифровая трансформация, государство и большие данные, опыт работы с большими данными

Заявление о конфликте интересов: Авторы заявляют об отсутствии конфликта интересов.

История статьи:

Поступила в редакцию: 04.02.2022. Принята к публикации: 30.03.2022.

Для цитирования:

Кузора С.С., Натаров И.П. Цифровая трансформация и большие данные // Вестник Российского университета дружбы народов. Серия: Государственное и муниципальное управление. 2022. Т. 9. № 2. С. 150–161. <https://doi.org/10.22363/2312-8313-2022-9-2-150-161>



Digital Transformation and Big Data

Stanislav S. Kuzora¹  , Ivan P. Natarov² 

¹Far Eastern Federal University,

8 Sukhanova str., Vladivostok, Russian Federation, 690091

²Ministry of Economic Development of the Primorsky Region,

22 Svetlanskaya str., Vladivostok, Russian Federation, 690110

 webkuzora@mail.ru

Abstract. Today, the Russian Federation pays special attention to the digital transformation of the country as a whole and shows a separate interest in big data in particular, which is confirmed by a number of analyzed documents and the current situation in the subject area. The research work proposed for reading is devoted to the analysis of big data, one of the areas of use of which is the state, or rather the work of federal and regional executive authorities. The main value of big data from the position of the state is seen in the possibility of working on volumes of heterogeneous information in order to increase efficiency in making managerial decisions on a wide range of issues. The main purpose of the study is to disseminate the experience of big data analysis, which is used in the work of the Ministry of Economic Development of Primorsky Krai. To do this, it is proposed to turn to the theoretical aspects of big data, find out the main chain of work on them, and also pay attention to domestic and foreign experience in using data in some areas of knowledge. The practical part of the study is a description of the experience of the Ministry of Economic Development of the PC, which is directly related to the digital development of the region and carries out a certain list of works with heterogeneous data. As an example, attention is drawn to a proven approach to the analysis of a large set of open data characterizing the work and development of small and medium-sized businesses in the Primorsky Territory, posted on the website of the Federal Tax Service of the Russian Federation. In conclusion, it is emphasized that the approach proposed in the study for working with data can be adapted to similar needs in other federal and regional authorities

Keywords: Big Data, data analysis, digital transformation, government and Big Data, Big Data practice

Conflicts of interest: The authors declared no conflicts of interest.

Article history:

The article was submitted on 04.04.2022. The article was accepted on 30.03.2022.

For citation:

Kuzora S.S., Natarov I.P. Digital Transformation and Big Data. *RUDN Journal of Public Administration*. 2022;9(2):150–161. <https://doi.org/10.22363/2312-8313-2022-9-2-150-161>

Введение

Перед тем как речь пойдет о таком явлении, как большие данные, обратим внимание на укрупненную область знаний «Наука о данных», представляющую собой междисциплинарную область, в основе которой лежат информатика и математика, применимая для интеллектуального анализа данных, машинного обучения и работы с большими данными. Наука о данных является как академической, так и практической дисциплиной межотраслевой сферы деятельности человека, направленная на изучение цифровых данных с целью извлечения полезной информации.

В теоретической части данной работы предлагается сосредоточиться на больших данных, а именно на основных этапах работы над ними: гене-

рация, сбор, хранение и анализ. Вместе с этим будет рассмотрено общее положение дел относительно предметной области на основе проведенного анализа. В завершении раздела делается переход от общего к частному, от использования больших данных во многих сферах деятельности к применению конкретного набора данных в определенной области. Надеемся, что такой способ изложения информации позволит читателю составить общее представление о больших данных и обратить внимание на частный случай их использования.

Теоретические аспекты больших данных

Обобщенно говоря, термин «Большие данные» используется для обозначения структурированных, полуструктурированных и неструктурированных объемов данных широкого многообразия [1; 2]. Существующее многообразие зависит от области знаний. Согласно фильтрам библиографической базы Scopus, большие данные охватывают 27 отраслей знаний. Первые пять направлений, к которым относятся компьютерные технологии, инжиниринг, математика, социальные науки и наука о принятии решений, в наибольшей степени сопряжены с большими данными. Отсюда может следовать, что для работы над данными обозначенных направлений могут использоваться ряд некоторых методов и инструментов [3], в то время как другие направления, менее связанные с большим объемом данных, предусматривают применение иных средств обработки информации.

Другой особенностью больших данных является набор признаков «трех жV» (объем, скорость, разнообразие), который получил свое развитие до пяти, в некоторых случаях до семи признаков [4]. Указанные признаки подчеркивают важность не только объема больших данных, но и другие существенные характеристики при работе с определенными наборами данных. Важно отметить, что объемы данных быстро изменяются, растут с течением времени, и работа над ними меняется с развитием и появлением новых технологий.

На сегодняшний день, как правило, большие данные измеряются в терабайтах, петабайтах, эксабайтах и зеттабайтах. По оценкам McKinsey & Company, к 2009 г. многие сектора экономики США имели в среднем не менее 200 терабайт хранимых данных (вдвое больше, чем хранилище данных американской розничной сети Walmart в 1999 г.) на компанию с более чем 1000 сотрудников [5].

Сегодня у общества изменилось отношение к данным по мере развития мультимедиа, социальных сетей, интернета вещей [6]. Теперь данные стали одним из важных производственных факторов для предприятий и бизнеса [7; 8]. В последнее время многие правительственные органы стран мира подключились к исследованиям и практическому применению больших данных [9]. Они также могут быть использованы одним челове-

ком в процессе интеллектуального анализа и аналитической обработки для принятия решений.

Помимо прослеживаемых преимуществ больших данных работа над ними сопровождается рядом проблем. Например, одной из основных является сложность в обнаружении наборов данных с большим масштабом, различными типами и быстрой генерацией с последующим извлечением ценной информации [10]. Другим примером сложности можно назвать непрерывное развитие инструментов обработки данных и появление новых технологий, предназначенных для высокоскоростного сбора, хранения и анализа, что приводит к дополнительным финансовым, иным затратам конечного пользователя [11].

После предварительного ознакомления с большими данными предлагается обратить внимание на основную цепочку работ над ними, состоящую из четырех этапов.

1) *Генерация* — первый шаг по работе с данными, нацелен на поиск информации в виде записей, файлов, сообщений в интересующей области. Такие данные имеют разные источники, формат, объем и другие отличительные характеристики. По отдельности они могут не представлять никакой ценности, однако их совокупность позволяет идентифицировать полезную информацию, на основе которой можно получить ответы на интересующие вопросы, отследить разного рода закономерности, узнать привычки людей, получить другого рода информацию.

2) *Сбор* — это следующий шаг по работе с данными, включающий процесс передачи и предобработки. Во время сбора данных с помощью определенных средств/инструментов накапливается огромный объем необработанной информации, которая может быть избыточной или бесполезной. После сбора необработанных данных происходит передача информации посредством специальных приложений в соответствующее хранилище. Для уменьшения информационной избыточности осуществляется предварительная обработка средствами технологии сжатия данных и участием специалиста, что позволяет эффективно хранить и использовать данные.

3) *Хранение* — процесс управления предобработанными данными, обеспечивающий надежность и доступность информации. Устройствами хранения данных являются системы управления баз данных (СУБД) реляционного типа (например PostgreSQL), NoSQL базы данных и т.д. Такие программные системы часто используются для широкого класса задач хранения данных, на основе которых можно разрабатывать уникальные приложения с целью постоянного взаимодействия с интересующими наборами данных. С повышенным ростом мировых объемов данных компании-поставщики вышеуказанных СУБД на постоянной основе увеличивают максимально допустимую емкость хранилищ, что указывает на высокую потребность в дополнительном пространстве хранения данных со стороны пользователей.

4) *Анализ* — является наиболее важным звеном при работе с большими данными, так как его результат призван дать некоторую ценность конечному

пользователю. До появления современных способов анализа данных применялись статистические методы анализа разнородных массивов данных. Наиболее распространенными являются регрессионный, корреляционный, кластерный анализ данных, которые посредством цифровизации стали применимы за счет программирования. Это привело к появлению отдельных библиотек, фреймворков в определенных языках программирования, например таких как *Rayton*, которые сочетают в себе знания статистики и информатики. Исходя из вышеперечисленного, такие возможности могут считаться современным способом анализа больших данных.

В качестве резюме по данному разделу следует отметить, что анализ больших данных играет огромную роль в понимании потребностей клиентов бизнеса, прогнозировании рыночных тенденций крупных организаций, разработке стратегий развития многих стран мира. Как может быть заметно, анализ больших данных достаточно трудоемкий процесс, который требует особых знаний и усилий для практического применения. В следующем разделе предлагается обратить внимание на роль больших данных в контексте цифровой трансформации Российской Федерации, после чего рассмотреть один из возможных способов анализа данных в процессе работы регионального органа исполнительной власти.

Цифровое государственное управление

Особое внимание в отношении развития цифровой экономики Российской Федерации на законодательном уровне уделяется с момента формирования национальной программы «Цифровая экономика Российской Федерации», утвержденная протоколом заседания президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам от 4 июня 2019 г. № 7.

В рамках национальных проектов России, сформированных Правительством Российской Федерации, федеральными и региональными органами исполнительной власти были инициированы такие проекты, как: цифровое государственное управление, кадры для цифровой экономики, цифровые технологии и др. [12–14]. Данные проекты направлены на ускоренное внедрение цифровых технологий с целью создания условий для высокотехнологичного бизнеса, повышения конкурентоспособности страны, укрепления национальной безопасности и повышения качества жизни граждан.

На основании проанализированных документов, утвержденных стратегий, запланированных проектов заметно, что большое внимание уделяется цифровым данным, которые охватывают многие аспекты социально-экономической деятельности в области государственного управления, образования, промышленности и бизнеса [15]. Несмотря на достаточно широкий перечень вышеуказанных направлений, предполагаемых быть основой для цифровой трансформации России, в рамках данной работы предлагается сфокусировать внимание на таком направлении, как «Большие данные», а именно на исполь-

зовании больших данных в деятельности Министерства экономического развития Приморского края.

Согласно нормативно-правовой базе, использование больших данных, а именно извлечение из них нужной информации, преимущественно направленно на совершение нормативной документации в государственном секторе, разработку новых образовательных программ, формирование соответствующих компетенций при работе с данными, создание технологических заделов для реализации новых проектов. Также отмечается, что для качественного управления цифровым развитием большие данные могут быть использованы как инструмент повышения прозрачности деятельности всех заинтересованных сторон.

Обратим внимание на частный случай необходимости использования больших данных в рамках цифровой трансформации отдельных отраслей экономики Приморского края. Основным документом преобразования является стратегия, цель которой заключается в повышении цифровой зрелости социально-экономической сферы по определенным направлениям, в том числе государственного управления. Со стороны государственного управления в стратегии обозначен ряд конкретных задач, выполнение которых с трудом представляется возможным без применения новых технологий и работы с большими данными.

Более того, в документе особая роль отводится проблемам и вызовам, с которыми придется столкнуться в процессе цифровой трансформации, где использование больших данных связано, например, с пунктом 4.5.8 «Отсутствие единого информационного пространства хранения, консолидации и обработки электронных данных для обеспечения своевременности и достоверности предоставления информации в органах исполнительной власти».

Также в стратегии перечисляется перечень запланированных проектов, требуемых ресурсов и установленных показателей для обеспечения реализации цифровой трансформации края. Ответственным органом исполнительной власти относительно цифрового государственного управления является Министерство цифрового развития и связи Приморского края. Однако стратегические ориентиры в преобразовании цифровой среды устанавливают и другие субъекты государственного управления, например, Министерство экономического развития Приморского края.

На сегодняшний день упомянутыми органами сформированы планы работы и мероприятия, включающие применение информационных технологий, в том числе новых методов и средств при работе с разными наборами данных. В виду достаточно раннего периода цифрового преобразования, особенно в области государственного управления, что подтверждается нормативно-правовой документацией и отсутствием реализации завершенных проектов в публичном информационном пространстве, работа с большими данными находится на начальной стадии формирования. Несмотря на небольшой срок развития данного направления, пример по работе с определенными наборами данных уже существует.

Опыт работы с большими данными

Как было упомянуто ранее, Минэкономразвития ПК имеет непосредственное отношение к цифровому развитию региона и осуществляет определенный перечень работ с данными. В этой связи предлагается обратить внимание на апробированный подход по анализу большого набора открытых данных, характеризующих работу и развитие малого и среднего предпринимательства в Приморском крае, размещенных на сайте Федеральной налоговой службы Российской Федерации.

Для удобства ознакомления с материалом структура раздела имеет следующий вид: обозначение актуальности, постановка цели и задач анализа, описание цепочки работ над данными, подведение итогов и рекомендации.

Актуальность. Одной из задач Министерства экономического развития Приморского края является анализ социально-экономической обстановки в регионе, формирование региональных механизмов развития, принятие взвешенных и обоснованных решений, направленных на развитие края. Сегодня в государственном секторе, различного рода информационными системами, ежегодно генерируются петабайты данных, которые зачастую не используются государственными служащими, а только размещаются в открытых источниках для внешних пользователей.

В качестве примера можно выделить сферу малого и среднего предпринимательства. Ежемесячно Федеральной налоговой службой (далее — ФНС) на основании Единого реестра субъектов малого и среднего предпринимательства и иных источников данных формируются гигабайты информации о развитии предпринимательства как в регионах, так и в Российской Федерации. В виду специфичной структуры данных их анализ усложняется и требует технических навыков для обработки и приведения в понятный вид.

Цель и задачи. Целью работы с точки зрения сформулированной актуальности является получение обобщающих показателей и выявление уровня налоговой платежеспособности малого и среднего бизнеса на территориях муниципальных образований Приморского края. Для достижения данной цели необходимо решение следующих задач: 1) сбор и предобработка данных; 2) формирование набора данных; 3) проведение статистического анализа данных.

Работа над данными.

Первый этап (сбор и предобработка данных). С официального сайта ФНС осуществляется выгрузка следующих данных, характеризующих развитие малого и среднего предпринимательства Приморского края:

1. Данные Единого реестра субъектов малого и среднего предпринимательства.
2. Сведения о суммах доходов и расходов по данным бухгалтерской (финансовой) отчетности организации за год.
3. Сведения о специальных налоговых режимах, применяемых налогоплательщиками.
4. Сведения о суммах недоимки и задолженности по пеням и штрафам.

5. Сведения об уплаченных организацией в календарном году суммах налогов и сборов.

В результате выгружено 54 481 файл, из которых:

- 6 211 файлов приходятся на дынные Единого реестра субъектов малого и среднего предпринимательства
- 12 063 файла содержат сведения о доходах и расходах;
- 12 067 файлов содержат сведения о специальных налоговых режимах;
- 12 056 файлов содержат сведения о суммах недоимки;
- 12 084 файла содержат сведения об уплаченных налогах.

Все данные имеют формат XML, что приводит к невозможности последующего анализа, и как следствие проводится преобработка данных языками программирования.

Второй этап (формирование набора данных). С помощью языка программирования Python были написаны скрипты по обработке полученных данных с последующим их приведением к формату CSV (Comma-Separated Values), удобного для использования библиотеками Python (Pandas) в табличном представлении.

В результате получены следующие наборы данных:

- reestr_msp_2019.csv;
- company_revenue_2019.csv;
- company_tax_regime_2019.csv;
- msp_arrears_2019.csv;
- msp_tax_2019.csv.

Далее полученные данные были объединены в итоговый набор данных (data_msp.csv) для проведения последующего статистического анализа. После объединения итоговый набор составляет 38 366 строк и 53 признака, где каждая строка характеризует одно юридическое лицо, имеющее признак субъекта малого и среднего предпринимательства, а признаки (колонки) — это характеристики данного субъекта, например, вид осуществляемой деятельности, категория бизнеса (микро, малый или средний), применяемая система налогообложения, выручка, уплаченные налоги, налоговая задолженность и т.д.

Третий этап (проведение статистического анализа данных). Заключительным этапом работы над данными является статистический анализ, который включает проведение исследовательского, корреляционного анализа и тестирование гипотез. Во избежание чрезмерного описания данного этапа предлагается кратко пояснить его составные элементы.

Исследовательский анализ данных (EDA — exploratory data analysis) предусматривает:

- описание данных (изучение размерности данных, сколько строк и столбцов в данных, описание типов данных, изучение мер центральной тенденции и распределения данных);
- изменение типов данных (преобразование не числовых значений в категориальный вид, приведение даты из строкового значение в формат даты и т.д.);

- описание пропусков (выявление процента пропущенных значений в данных и последующее их удаление или восстановление);
- описание дубликатов (поиск дублирующей информации в данных);
- работа с категориальными признаками (проверка орфографии, стандартизация наименований при необходимости и т.д.);
- вывод (результат расчетов, отображение графиков, построение корреляционной матрицы и т.д).

Для последующего анализа были сформулированы две гипотезы.

Гипотеза № 1. Муниципальные образования отличаются по уровню налоговой платежеспособности у субъектов малого и среднего предпринимательства Приморского края.

Гипотеза № 2. Существует группа муниципальных образований Приморского края, имеющая наиболее низкий уровень налоговой платежеспособности у субъектов малого и среднего предпринимательства Приморского края.

Тестирование гипотез проводилось с помощью критерия Краскела-Уоллиса и в результате получены следующие выводы:

- муниципальные образования Приморского края отличаются по уровню налоговой платежеспособности субъектов малого и среднего предпринимательства;
- существуют группы муниципальных образований Приморского края, имеющие наиболее низкий уровень налоговой платежеспособности у субъектов малого и среднего предпринимательства Приморского края.

Ссылка на полный аналитический отчет по проделанной работе представлена в списке использованных источников [16].

Итоги и рекомендации. По результатам проделанной работы цель по получению обобщающих показателей и выявлению уровня налоговой платежеспособности у субъектов малого и среднего предпринимательства Приморского края достигнута. Установлено, что муниципальные образования отличаются по уровню налоговой платежеспособности и существует группа муниципальных образований, имеющая наиболее низкий уровень налоговой платежеспособности.

В качестве рекомендаций можно отметить, что организациям инфраструктуры поддержки субъектов малого и среднего предпринимательства Приморского края, при формировании годового плана обучающих мероприятий бизнеса по налоговой грамотности, необходимо обратить внимание на муниципалитеты с низким уровнем налоговой платежности.

Другая рекомендация может быть адресована контрольно-надзорным органам, наделенным полномочиями в области налогового контроля, в целях повышения уровня поступления налогов в консолидированный бюджет Приморского края необходимо усилить налоговый контроль на территориях муниципалитетов.

Заключение

Подводя итог исследованиям, предлагается остановиться на ключевых выводах и особенностях проделанной работы:

- большие данные являются неотъемлемой частью цифровой экономики Российской Федерации, грамотное использование которых способно привести к достижению национальных целей и решению стратегических задач;
- реализация таких федеральных и региональных проектов, как «Цифровое государственное управление», а также стратегий в области цифровой трансформации неразрывно связана с работой над разнородными данными;
- знакомство с подходом по работе с определенными наборами данных при участии Министерства экономического развития Приморского края показало один из возможных вариантов анализа неструктурированных данных;
- предлагаемый подход имеет потенциал адаптации под работу с разными наборами данных в сферах деятельности как Минэкономразвития ПК, так и в других федеральных и региональных органах исполнительной власти Российской Федерации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Chen M., Mao S., Liu Y.* Big Data: A Survey // *Mobile Netw Appl.* 2014. P. 171–209. DOI: <https://doi.org/10.1007/s11036-013-0489-0>
2. *Longbing C.* Data Science: A Comprehensive Overview // *ACM Comput. Surv.* 2017. Vol. 50. Issue 3. Article 43. P. 42. DOI: <https://doi.org/10.1145/3076253>
3. *Звягин Л.С.* Цифровые тренды в анализе данных и мягких измерениях как концептуальная основа развития прикладных наук // *Мягкие измерения и вычисления.* 2020. Т. 37. № 12. С. 45–62.
4. *Chen J., Jiang Q., Wang Y., Tang J.* Study of Data Analysis Model Based on Big Data Technology // 2016 IEEE International Conference on Big Data Analysis (ICBDA). 2016. P. 1–6. DOI: <https://doi.org/10.1109/ICBDA.2016.7509810>
5. *Big Data: The Next Frontier for Innovation, Competition, and Productivity / McKinsey Global Institute.* URL: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#>. Дата обращения: 23.01.2022.
6. *Платонова С.И.* Большие данные: создание вызовов и возможностей в социальных науках // *Манускрипт.* 2020. Том 13. Выпуск 4. DOI: <https://doi.org/10.30853/manuscript.2020.4.24>
7. *Черняк Л.* Большие Данные — новая теория и практика // *Открытые системы. СУБД.* 2011. № 10. С. 18–25.
8. *Черняк Л.* Свежий взгляд на Большие Данные // *Открытые системы. СУБД.* 2013. № 7. С. 48–51.
9. *Волков Д.В.* В поисках сокровищ // *Открытые системы. СУБД.* 2014. № 1. С. 1.
10. *Sagiroglu S., Sinanc D.* Big Data: A Review // 2013 International Conference on Collaboration Technologies and Systems (CTS). 2013. P. 42–47. DOI: <https://doi.org/10.1109/CTS.2013.6567202>.
11. *Philip Chen C.L., Chun-Yang Z.* Data-intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data // *Information Sciences.* 2014. Vol. 275. P. 314–347. DOI: <https://doi.org/10.1016/j.ins.2014.01.015>

12. Региональные проекты цифровой экономики Приморского края. URL: <https://digital.primorsky.ru/>. Дата обращения: 23.01.2022.
13. Указ Президента РФ от 09.05.2017 N 203 «О Стратегии развития информационного общества в Российской Федерации на 2017 — 2030 годы». URL: <https://base.garant.ru/71670570/>. Дата обращения: 23.01.2022.
14. Указ Президента РФ от 21.07.2020 N 474 «О национальных целях развития Российской Федерации на период до 2030 года». URL: <http://publication.pravo.gov.ru/Document/View/0001202007210012>. Дата обращения: 23.01.2022.
15. Стратегия в области цифровой трансформации отраслей экономики, социальной сферы и государственного управления Приморского края. URL: https://www.tadviser.ru/images/7/7c/Стратегия_Приморский_край.pdf. Дата обращения: 23.01.2022
16. Проект по анализу налоговой платежеспособности малого и среднего бизнеса на территориях муниципальных образований Приморского края. URL: https://github.com/ivannatarov/Data_analysis_msp. Дата обращения: 23.01.2022

REFERENCES

1. Chen M., Mao S., Liu Y. Big Data: A Survey. *Mobile Netw Appl*, 2014: 171–209. DOI: 10.1007/s11036-013-0489-0
2. Longbing C. Data Science: A Comprehensive Overview. *ACM Comput. Surv.* 2017;50(3):42. DOI: <https://doi.org/10.1145/3076253>
3. Zvyagin L.S. Cifrovye trendy v analize dannyh i mjagkih izmerenijah kak konceptual'naja osnova razvitija prikladnyh nauk [Mathematical Algorithms of Game Theory as an Applied Tool for Making Effective Financial and Economic Decisions]. *Myagkiye izmereniya i vychisleniya*. 2020;37(12):45–62 (In Russ.).
4. Chen J., Jiang Q., Wang Y., Tang J. Study of Data Analysis Model Based on Big Data Technology. *2016 IEEE International Conference on Big Data Analysis (ICBDA)*; 2016: 1–6. DOI: <https://doi.org/10.1109/ICBDA.2016.7509810>
5. Big Data: The Next Frontier for Innovation, Competition, and Productivity. *McKinsey Global Institute*. URL: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#>. Accessed: 23.01.2022
6. Platonova S.I. Big Data: Challenges and Opportunities in Social Sciences. *Manuskript*. 2020; 13(4). DOI: <https://doi.org/10.30853/manuscript.2020.4.24> (In Russ.).
7. Chernyak L. Bol'shiye Danyye — novaya teoriya i praktika [Big Data — New Theory and Practice]. *Otkrytyye sistemy. SUBD*. 2011;10:18–25 (In Russ.).
8. Chernyak L. Svezhiy vzglyad na Bol'shiye Danyye [A New Look at Big Data]. *Otkrytyye sistemy. SUBD*. 2013;7:48–51 (In Russ.).
9. Volkov D.V. V poiskakh sokrovishch [Looking for Treasures]. *Otkrytyye sistemy. SUBD*. 2014;1:1 (In Russ.).
10. Sagioglu S., Sinanc D. Big Data: A Review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*; 2013: 42–47. DOI: <https://doi.org/10.1109/CTS.2013.6567202>
11. Philip Chen C.L., Chun-Yang Z. Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data. *Information Sciences*. 2014;275:314–347. DOI: <https://doi.org/10.1016/j.ins.2014.01.015>
12. *Regional'nye proekty cifrovoj jekonomiki Primorskogo kraja* [Regional Projects of the Digital Economy of Primorsky Region]. URL: <https://digital.primorsky.ru/>. Accessed: 23.01.2022 (In Russ.).
13. *Ukaz Prezidenta RF ot 09.05.2017 N 203 «O Strategii razvitija informacionnogo obshhestva v Rossijskoj Federacii na 2017 — 2030 gody»* [Decree of the President of the Russian Federation of May 9, 2017 # 203 “On the Strategy for the Development of the Information Society in the Russian Federation for 2017 — 2030”]. URL: <https://base.garant.ru/71670570/>. Accessed: 23.01.2022 (In Russ.).
14. *Ukaz Prezidenta RF ot 21.07.2020 N 474 «O nacional'nyh celjah razvitija Rossijskoj Federacii na period do 2030 goda»* [Decree of the President of the Russian Federation of July

- 21, 2020 # 474 “On the National Development Goals of the Russian Federation for the Period up to 2030”]. URL: <http://publication.pravo.gov.ru/Document/View/0001202007210012>. Accessed: 23.01.2022 (In Russ.).
15. *Strategija v oblasti cifrovoj transformacii otraslej jekonomiki, social'noj sfery i gosudarstvennogo upravljenja Primorskogo kraja* [Strategy in the Field of Digital Transformation of Sectors of the Economy, Social Sphere and Public Administration of Primorsky Region]. URL: https://www.tadviser.ru/images/7/7c/Стратегия_Приморский_край.pdf. Accessed: 23.01.2022 (In Russ.).
16. *Proekt po analizu nalogovoj platezhеспособности malogo i srednego biznesa na territorijah municipal'nyh obrazovaniy Primorskogo kraja* [Project to Analyze the Tax Solvency of Small and Medium-sized Businesses in the Territories of Municipalities of Primorsky Region]. URL: https://github.com/ivannatarov/Data_analysis_msp. Accessed: 23.01.2022 (In Russ.).

Информация об авторах:

Кузора Станислав Сергеевич — ассистент Департамента инноваций Политехнического института Дальневосточного федерального университета, ORCID: 0000-0003-2309-2035, e-mail: kuzora_ss@dvfu.ru

Натаров Иван Петрович — аналитик данных Министерства экономического развития Приморского края, ORCID: 0000-0002-6550-6833, e-mail: natarov_ivan@bk.ru

Information about the authors:

Stanislav S. Kuzora — Assistant of the Department of Innovations, Polytechnic Institute, Far Eastern Federal University, Russian Federation, ORCID: 0000-0003-2309-2035, e-mail: kuzora_ss@dvfu.ru

Ivan P. Natarov — Data Analyst, Ministry of Economic Development of the Primorsky Region, Russian Federation, ORCID: 0000-0002-6550-6833, e-mail: natarov_ivan@bk.ru