Public Organization Review
https://doi.org/10.1007/s11115-022-00694-x

# Digital Transformation: Exploring big data Governance in Public Administration

**Alexander Yukhno**[1]

Accepted: 8 December 2022


**Abstract**
As economies become increasingly data-driven, big data technologies and software products are turning into key tools for managing technological processes in real time for more efficient delivery of public services to citizens. The change in the interaction between the state and society implies the creation of a unified state digital ecosystem, centered around big data. Such paradigm calls for rethinking of public administration principles ensuring the transition from an electronic state to a digital one. As a result, the likelihood of creating values that meet the shifting expectations of citizens in relation to public services increases.

**Keywords** Artificial intelligence · Big data · Digital transformation · Public administration

## Introduction

Nowadays big data technologies and software products are turning into key tools for managing digital transformation. The development of the state and society interaction implies the creation of a unified state digital ecosystem with dig data placed in the center. The global big data market is projected to grow to US$ 103 billion by 2027, more than double the market size in 2018. At the same time software will become the largest big data market segment by 2027 with a share of 45% (Mlitz, 2021).

Investments in big data, artificial intelligence and machine learning technologies (Yukhno, 2019) are essential for the development of this area. Its volume depends on

✉ Alexander Yukhno
alexander_yukhno@mail.ru

1 Institute of Public Administration and Civil Service of the Russian Presidential Academy of National Economy and Public Administration, 76, Prospect Vernadskogo, 119454 Moscow, Russia

Published online: 13 December 2022

the industry and market segment. For example, the 2021 Global Data survey shows that for the pharmaceutical industry, big data (51%) and artificial intelligence (51%), along with APIs and digital platforms (39%), will be the most important investments in the next two years (*Big data and AI are likely to lead emerging technology investments: Survey*, 2021). Surveys show that most of these investments spent on the budget for ICT solutions (55%), training employees (48%), hiring new employees (30%) and paying for external technical consultants (26%). Financing of such investments mainly comes from the IT budget (46%), implying that big data is still considered a technical issue (*Big data is a job engine*, n.d).

The COVID-19 pandemic has contributed to the implementation into practice the tools and processes for working with big data, which reflects the key technological trends of state and corporate development and have the greatest potential impact on the competitive advantages of states and organizations in the coming years.

## Methodology

The current study presents a comprehensive analysis of big data governance in public administration.

The author applies empirical and comparative analysis, expert assessments, synthesis, deduction, and induction to draw the conclusions.

Big data is becoming an increasingly important intangible asset for states and organizations that affects the implementation of their strategies and competitive advantages in the market. Thus, the key technological trends in a state and corporate development concerning data-driven public administration in the post-Covid period were examined.In this context the following aspects of big data application were studied: the main features of big data and reasons for introducing it to public administration, the critical challenges to its efficient application and a data-driven public sector. Based on the results obtained, the need to create a unified state digital ecosystem with data and algorithms in the center is logically justified.

The article begins with a brief overview section. The next part presents the analysis and findings about efficient big data application. It is followed by discussion on a path to efficient data-driven public administration and its implications for states and organizations. The article is concluded with final remarks and discissions for further research.

## Big Data As a New Effective Tool For States and Organizations

Altough big data existed in the past, the widespread adoption of the term is credited to C.Lynch, the editor of the Journal Nature, who used it in 2008 in an issue dedicated to the explosive growth of the world's information (Lynch, 2008). Gartner understands big data as high-volume, -velocity and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Sicular, 2013). ISO/IEC 20546:2019 (Information technology — Big data — Overview and vocabulary) defines big data as extensive datasets — primarily in the data characteristics of volume, variety, velocity, and/or variability — that require a scalable technology for efficient storage, manipulation, management, and analysis (*GOST R ISO/IEC 20546–2019. Information technology. Big data. Overview and vocabulary*, 2020). In essence, the term describes the exponentially increasing volume and variety of data belonging to organizations and states that cannot be processed using conventional tools (Bennett, 2017, p. 462).

The main sources of big data include mobile and social platforms, streaming data that comes from the Internet of Things (IoT), public data from open sources, cloud data, databases (Oracle, SQL, Amazon Simple, etc.), ERP systems, data warehouses, etc. (*Big data: what is it and why it matters*, n.d; Joshi, 2017). In most cases, data is generated by an extremely heterogeneous user base and social and organizational mechanisms that differ significantly from the technical models that have traditionally been used in economic life. The situation is further complicated by the fact that users of the Internet and social networks often do not belong to the organizations or networks to which they transmit data (Constatiou and Kallinikos, 2015, pp.4–5). In this regard, the issue of ensuring its confidentiality becomes extremely relevant.

In recent years, researchers have developed various features of big data that characterize this term and reveal its content[1]. The author proposes to highlight the key ones among them:

1. Volume - the amount of big data created, collected, copied, and consumed around the world is growing exponentially and will reach 180 zettabytes by 2025 (Holst, 2021).
2. Velocity - big data is created at a tremendous speed and requires real time processing, storage, and analysis.
3. Variety - big data is formed from various types of data in a wide variety of formats (structured, semi-structured and unstructured data, text data, speech data, video data, etc.).
4. Veracity - big data must be characterized by accuracy, truthfulness, and reliability, otherwise they can lead to incorrect decisions.
5. Variability - understanding the meaning of big data requires consideration of the chosen context.
6. Visualization - for greater accessibility of perception of big data, its visualization is required.
7. Value - the analysis of big data allows you to derive value to solve the problems of specific users (Ishwarappa, 2015; Sicular, 2013; Rialti et al., 2016).

Nowadays big data is used in almost all sectors of the economy and is finding new areas of application, such as identifying consumer buying habits, personalized marketing, health monitoring using data from wearable devices, road mapping for autonomous vehicles, customized health plans for cancer patients, traffic analysis to reduce

[1] In practice, there are more than 50 characteristics of big data. For example, Dhamodharavadhani, S., Gowri, R., Rathipriya, R. (2018, March). *Conference: First International Conference on Computer Vision,Networks and Informatics.* Unlock Different V's of Big Data for Analytics. *International Journal of Computer Sciences and Engineering*, 06 (04), 183–190.

congestion in cities, production optimization, real-time data monitoring and cybersecurity protocols, etc.

T.H. Davenport and J. Dyché highlight the following among the main reasons for using big data technologies: cost and task execution time reduction, development of new products and offerings, and support of internal business decisions (Davenport and Dyché, 2013, pp. 3–8). This position could be complemented by such benefits as identifying potential risks that can improve the quality and effectiveness of the risk management system, the ability to create targeted products and services for a specific market, retaining and attracting new customers (Mills, 2019).

The COVID-19 has only accelerated the indicated trends, demonstrating the key role of big data in promptly responding to emerging challenges and making decisions that affect all or most people in states and organizations[2]. It has rapidly influenced many aspects of governments and organizations activities, significantly increasing the role of artificial intelligence and machine learning in big data analytics in the public and private sectors. Due to the need to collect, segment and promptly interpret huge amounts of data related to the spread of the virus, the important role of artificial intelligence algorithms in this process was revealed (Marr, 2020). In addition, due to the use of artificial intelligence, states and organizations received a functional tool which can predict, detect, diagnose and treate various viruses for the coming years in real-time mode (Ibid). At the same time, states and organizations are facing problems in the field of data protection (issues of confidentiality, security, data leakage, etc.), which have underlined the need for new approaches to regulating this area of our life.

## A Path to Efficient Big Data Application

Big data usability depends on two key factors, which are the main goals of the applied business processes: the ability to create value from it and the quality of such data (Hilb, 2019, p. 60). V.D. Milovidov notes that the presence of a large amount of data and technologies for its processing does not have “any pronounced correlation with the economic results of economic activity.“ As a confirmation of this thesis, the author points out that “the total volume of accumulated data in American companies practically does not correlate with the indicators of their market capitalization” (Milovidov, 2019, p. 288). In reality, the presence of a large amount of generated data does not always allow to determine at first glance potential options for extracting value from it and using it for solving practical problems. On average, from 60 to 73% of all data in an organization is not used for analytics, which considering the specifics of data-driven economy may lead to a decrease in competitiveness (Gualtieri, 2016). Moreover, the intended effects of data analysis cannot be verified retrospectively, and the transparency of the data can cause controversy among the participants in the digital ecosystem. The Organization for Economic Cooperation and Development (OECD),

[2] For example, working with big data has enabled the FTS of Russia to effectively provide targeted support measures to citizens and organisations in a short period of time. Federal Tax Service (2020, November 12). *Na obshchestvennom sovete pri FNS Rossii obsudili rabotu Sluzhby vo vremya pandemii [The Public Council of the Russian FTC discussed its work during the pandemic]*. Retrieved May 21, 2022, from https://www.nalog.ru/rn77/news/activities_fts/10154705/.

while noting the inexhaustibility of data, emphasizes that it can be reused to open up significant growth opportunities or create benefits for the whole society in ways that could not have been foreseen when they were first created (*Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, 2019). With more data being produced today than ever before, the ability to use it becomes the key competitive advantage for states and organizations. Understanding and interpretation of the data can provide them with valuable information needed to improve its performance (forecasting demand, potential problems, etc.).

The issue of data quality is one of the key areas of activity of states and organizations, not just their ICT departments (Yukhno, 2021). During the coronavirus pandemic, the low quality of data did not allow the use of various artificial intelligence models for diagnosing the disease and predicting risks for patients, as they were not suitable for clinical use. Thus, it becomes clear that poor data quality jeopardizes the achievement of state goals and largely depends on the quality of the process for data obtaining, as well as on the quality of the processes for its storing, managing, transferring, and presenting (*GOST R 54524–2011/ISO/TS 8000–100:2009. Data quality. Part 100. Master data. Exchange of characteristic data. Overview*, 2012). At the same time, according to GOST R 56214–2014 / ISO / TS 8000-1: 2011, data quality affects only those data that are involved in making any decision, while preventing the repetition of data defects and reducing unnecessary costs (*GOST R 56214–2014/ISO/TS 8000-1:2011. Data quality. Part 1. Overview*, 2015). In this regard, interaction in the field of data management should include stakeholders with a clear division of areas of responsibility and be directly linked to the development priorities of states and organizations (Goasduff, 2020).

A report by Precisely and Corinium Global Intelligence shares that data reliability implies three characteristics: accuracy, consistency, and context. The integration of cloud technologies helps improve the quality of collection and analysis of big data. However, 82% of data leaders find it difficult to provide a constant stream of reliable data that is suitable for making informed business decisions. In addition, 9 out of 10 managers face a shortage of workers with the right skills (*Data Integrity Trends: Chief Data Officer Perspectives in 2021*, 2021; *Precisely: 82% of data executives cite data quality as a barrier*, 2021).

> In the digital era, data quality needs to be directly linked to the KPIs of governments and organizations (Moore, 2018b). Data collection requires certain costs and clear regulations for assessing the use and profitability of acquiring various types of data for specific tasks. However, according to Gartner, about 60% of organizations do not even measure the annual financial costs (lost revenue opportunities, reduced organizational efficiency and productivity, etc. (Barrett, 2022) on low quality data (Moore, 2018b).
> Governments and organizations need to clearly understand the relationship between additional data, improving forecasting accuracy and creating added value (Heath, 2019, p. 46). It also means quantifying the financial benefits of big data, as well as the costs (and consequent savings) of investing in risk management in this area (Bennett, 2017). That said, according to a Gartner survey, 42% of data analytics leaders do not measure or track metrics in this

area. Those who do carry out such activities are mainly focused on achieving compliance goals (Goasduff, 2020). In this regard, it is recommended that states and organizations:

- establish a unified approach to data;
- measure the values of big data indicators;
- identify key data workers;
- optimize the cost of data quality management tools;
- set realistic time frames for deploying data quality tools (Moore, 2018b);
- implement a quality control system for big data;
- create a unified repository of big data and provide access to it for all employees in accordance with internal regulations;
- develop a corporate culture of working with big data;
- implement employee-friendly big data analytics tools.

Big data is becoming an increasingly important intangible asset (*The world's most valuable resource is no longer oil, but data: Regulating the Internet giants*, 2017) for states and organizations, contributing to the creation of added value in various sectors of the economy. Such approach contributes to digital transformation (Dremel et al., 2017, p. 97) and gaining competitive advantages in the market (Ibid) by predicting needs and responding to changes in the economy (*Strengthening Digital Government*, 2019, p. 3). In other words, big data needs to be viewed as an asset and both its value and the value it stores should be measured. In this regard, states and organizations need to implement measurable indicators that will allow to trace the relationship between data analytics and ongoing initiatives and projects to create value and achieve their goals and objectives. Specifically, they are invited to:

1. determine the key priorities of activities;
2. approve KPIs (financial, operational, reputation, etc.) in achieving goals and objectives using big data analytics;
3. determine the current level of quality of existing big data and its impact on their activities;
4. approve the target state of data quality to achieve the set goals (Moore, 2018a).

Currently the importance of solutions that allow monitoring of real time data quality metrics is growing. States and organizations are now using artificial intelligence and machine learning to support more complex data management tasks, including data cataloging, metadata managing, and displaying and anomalies detecting (Hunt, 2021). In addition, the use of artificial intelligence and machine learning contributes to significant improvements in the quality of the data itself as an intangible asset. At the same time, high-quality and compatible data is required, for the efficient operation of artificial intelligence (*Artificial Intilligence. Towards a choice of strategy*, 2019, p. 56). In this regard, states and organizations, first of all, need to systematically consider the issue of introducing strategies and tools for managing such data. It is noted that aligning strategies for big data, cloud computing and mobile technologies leads to a 53% economic growth in comparison to competitors that do not use these technologies (Columbus, 2015).

High data quality provides the ground for an effective data management structure, supports business initiatives, minimizes emerging risks, and ensures the reliability, completeness, uniqueness, and relevance of data itself. In practice, the main tools for improving the quality of data are as follows: adopting a data management strategy, data profiling, data checking as it is entered, data cleaning, data quality monitoring and controlling, building interaction between organizational units, making changes to the organizational structure (for example, the introduction of data manager position or the creation of a data quality control center) with the setting of appropriate business processes, roles, functionality, stakeholder responsibilities and realistic deadlines for the implementation of these tools in practice. At the same time, the widespread introduction of tools to improve data quality, as a rule, is undermined by their high cost (Moore, 2018b).

With the exponential scaling of the volume of big data moving between nations and organizations, value chains will become more complex and analytical tools more and more diversified. Thus, states and organizations need to carry out continuous and comprehensive work to systematize and control big data quality used to create in order to form an appropriate corporate culture of big data management for further transition to big data-based decision-making.

## Moving to a Data-Driven Public Sector

The OECD classifies a data-driven public sector as one of the six dimensions of the OECD Digital Government Framework (*The OECD Digital Government Policy Framework: Six dimensions of a Digital Government*, 2020, p. 6), highlighting that a mature digital state is a data-driven state (Ibid, p. 14) and calls for a shift from information-driven states to a data-driven public sector[3]. In real life, the implementation of big data technologies in public administration faces three main issues:

- technical and practical problems (insufficient amount of quality data, lack of common standards and level of interaction between various public and private ICT systems, etc.);
- limited resources and opportunities (insufficient funding for necessary research, low qualifications of civil servants and digital literacy in society, etc.);
- the presence of barriers in society and the state (institutional, legislative, cultural, etc.) (Ubaldi et al., 2019, p. 53).

To simplify and improve the abovementioned practices, the OECD recommends that states develop strategies to support data-driven public sector development (*Strengthening Digital Government*, 2019, p. 1). Typically, such a strategy focuses on three

[3] In practice, the terms “data” and “information” are often used interchangeably. However, it is useful to distinguish between them. Data consists of bits and bytes, which are created by computer systems. Information comes from data, after the data has been organised, analysed and presented in a particular context.

main areas: the needs of citizens and society, obtaining, analyzing, and making decisions based on big data, mobility, and speed of sharing big data and decision making (Akatkin et al., 2017, p. 19). In addition, the data-driven state focuses on its use to create social value by predicting and planning possible changes, improving the quality of public services, implementing public policy and responding to public requests, as well as assessing and monitoring its use (*The OECD Digital Government Policy Framework: Six dimensions of a Digital Government*, 2020, p. 15).

In practice, this means that the development of approaches in which new technologies and big data become instruments for achieving the state's goals and objectives. At the same time, the very transition to strategic management of big data at the state level, as a rule, presupposes the presence of three major elements:

- nationwide data architecture;
- digital transformation of public administration;
- human resources (*The state as a platform. People and technology*, 2019, p. 8–11).

In addition, the analysis shows that the shift to data-driven governance can be facilitated by a government strategy focusing on two main areas: leveraging existing data from different sources through aggregation and analysis and building real-time data exchange networks to deliver improved public service (*Real-Time, Data-Driven Government. Develop Forward-Thinking, Citizen-Centric Programs*, 2021, p. 9).

The change in the format of interaction between the state and society necessitates the creation of a unified state digital ecosystem with data and algorithms in the center. The currently implemented platform approach in public administration[4] will significantly increase the speed of provision of public services, and feedback from citizens in real time will significantly increase the requirements for the quality of their provision.

Such a paradigm of relations presupposes a rethinking of the principles of the public administration system, ultimately ensuring the transition from an electronic state to a digital one (*The OECD Digital Government Policy Framework: Six dimensions of a Digital Government*, 2020, p. 6). The goal of this process should be the transformation of all areas of the state's activities, including internal processes and standards for the provision of public services, the organizational structure of public administration, internal corporate culture, the necessary skills and competencies of civil servants, the format of interaction with citizens and receiving feedback, etc. As a result, the likelihood of creating values that meet the increased expectations and requirements of citizens regarding the results of their interaction with the state increases. Thus, in the digital era, the digital transformation of the state necessitates the adaptation and use of new technologies and big data in state activities in order to "reset" public administration.

[4] For example, Petrov, M., Burov, V., Shklyaruk, M., Sharov, A. (2018, April). *The State as a Platform: People and Technology*. Center for Strategic Research. Retrieved May 24, 2022, from (https://www.csr.ru/upload/iblock/313/3132b2de9ccef0db1eecd56071b98f5f.pdf).

In this regard, World Bank's "World Development Report 2021: Data for Better Lives" is proposing a new social contract to address data governance challenges that will:

- increase the use and reuse of data to extract more value;
- ensure more equitable access to the benefits that data creates;
- build trust by creating mechanisms to protect citizens from data risks;
- prepare the ground for the creation of an integrated national data system (*World Development Report 2021: Data for Better Lives*, 2021).

An integrated national data system includes three groups of actors: states and international organizations, individuals and civil society, and the private sector, which cooperate in an environment where data is safely produced, exchanged, and used. Despite the need for close interaction between the parties, such a system does not require an obligatory centralized management system and storage of all data in it (Ibid).

The introduction of the term "unified national data system" into scientific research reflects a new stage in the development of the digital environment. The system is built on the basis of four pillars (Table 1).

Both states and organizations that want to be effective during the period of active digital transformation of the economy, when moving to data-centric management, are recommended to take into account the above factors.

The priority of using big data for the state in the Russian Federation is consolidated in regulatory documents. The decree of the President of the Russian Federation of 05/09/2017 N 203 "On the Strategy for the Development of the Information Society in the Russian Federation for 2017–2030" defines the digital economy as "an economic activity in which the key production factor is digital data, processing of large volumes and the use of the analysis results of which, in comparison with traditional forms of management, can significantly increase the efficiency of various types of production, technologies, equipment, storage, sale, delivery of goods and services" (Strategy of the Information Society Development in the Russian Federation for 2017–2030, 2017). On June 3, 2019, by Order of the Government of the Russian Federation No. 1189-r, the Concept for the Creation and Operation of a National Data Management System (NDMS) and a roadmap for the creation of an NDMS for 2019–2021 were approved.

According to the Big Data Association, compared to 2019, the introduction of big data in the Russian Federation under the optimistic scenario can provide 1.8% of GDP growth (1.2% under the baseline scenario), as well as the growth of the big data market to 160 billion rubles (100 billion rubles - under the baseline scenario) by 2024. At the same time, in order to expand the possibilities of using big data, the Association proposes to overcome five groups of barriers (lack of specialists for the mass adoption of big data, limited infrastructure and data availability, difficulties in research and innovation in the field of big data, lack of large-scale implementation of big data in sectors of the economy) (*Strategy for the Big Data market development by 2024*, 2019).

Currently, in order to be competitive, both states and organizations need to use a wide range of internal and external information resources. This means increasing the

**Table 1** Four pillars of a unified national data system

| Pillar | Specifications |
|---|---|
| Infrastructure policy | provides equal access to data for states and people;<br>contributes to the development of internal data infrastructure, which allows for local storage, processing and exchange of data within the state. |
| Data Regulations | trust in data transactions is supported by strong laws and regulations;<br>factors facilitating the exchange of data are usually more developed for publicly available data, where government policy and legislation establish rules for access and exchange of such data, than for private data, where governments have more limited influence. |
| Economic policy measures | growing role of data in the business models of digital platforms is changing competition, trade and taxation in the real economy, posing risks for low- and middle-income states;<br>security measures and data processing tools developed by states will have an impact on the real economy (antitrust regulation, tax issues, trade in services based on data);<br>low- and middle-income states often lack the institutional capacity to manage the policy problems posed by data-driven economies. |
| Data Management Institutes | institutions to manage data have four main functions: strategic planning, setting rules and standards, compliance and enforcement, and gaining the knowledge and evidence needed to understand and address emerging issues;<br>a targeted, participatory approach to data management and oversight can help institutions keep pace with the ever-evolving data ecosystem and enhance their legitimacy, transparency and accountability (*World Development Report 2021: Data for Better Lives*, 2021). |

volume of different types of data by closer collaboration with different stakeholders (Judah, 2021). It seems that effective data management does not occur in isolation, but in close interaction with other participants of the digital ecosystem, who are collectively responsible for the final result. In this regard, it is necessary to work out in detail the issue of the availability of open and common tools (for example, open standards, APIs, algorithms) that can facilitate overall integration both inside and outside the organization (*The Path to Becoming a Data-Driven Public Sector*, 2019). Considering the problem from the state point of view, the establishment of partnerships outside the public sector will allow it to:

- take advantage of digital private sector solutions to improve, optimize and modernize public sector data infrastructure (e.g. cloud solutions);
- facilitate the publication of data produced by civil society organizations on government open data platforms or the publication of open government data on non-government data portals;

- support the exchange of data between multiple stakeholders from different sectors and strengthen the control and power of data owners over the exchange and use of their data for common policy objectives (Ibid).

For example, the practice of collecting big data is gaining popularity in order to maintain economic statistics used to monitor the economic situation in the state (in Indonesia, tourism statistics are monitored using data from cellular operators, Azerbaijan uses data from electronic databases of retail for statistical purposes, in Australia statistics on jobs and gross wages are analyzed using payroll data, etc.) (Bernal et al., 2021).

## Concluding Discussions

This paper aims to fill the knowledge gap around big data and data-driven public administration in the post-Covid period. The pandemic has intensified the key role of big data in responding to emerging challenges and decision-making that affects states and societies, significantly increasing the role of artificial intelligence and machine learning in big data analytics in the public and private sectors. The author states that the volume of analyzed big data is constantly growing and requires the creation and maintenance of an appropriate technological architecture and ICT infrastructure, as well as an increase in the scale of its storage, analysis, and transmission. At the same time, the fact of having big data does not automatically impy that it can be used to solve real problems. The main value of big data lies not in itself, but in the conclusions, products and services that appear as a result of its processing and analysis (Davenport and Dyché, 2013, p. 30). In fact, organizational productivity improvements and competitive advantage arise from analytic models that identify new opportunities and enable states and organizations to predict and optimize their performance (Barton and Court, 2012). The author stresses that data reliability implies such characteristics as accuracy, consistency, and context. The issue of data quality is one of the key areas of activity of states and organizations and needs to be directly linked to their KPIs. However. the widespread introduction of tools to improve data quality nowadays is constrained by their high cost.

The analysis reveals problems in the field of data protection. The use of big data leads to a reset of the risk management system both within the state and organizations, allowing it to be processed in real time to obtain more accurate results, improve the quality of work and feedback from citizens and counterparties. The main task in the field of big data is finding a balance between emerging risks and its value for the state and organizations in the context of the constant increase in volume and the development of new technologies that ensure its cheaper storage. In addition, the growth in big data can lead to an information gap between the strategic and operational levels of government and corporate governance. In this regard, one of the important consequences of the digital transformation of an organization is the need to build processes that could be managed in real time. Mechanisms for managing big data, on the one hand, have to protect it from factors that can destroy or limit its value, on the other hand, to exclude situations in which states and organizations will not be able to take of advantage of the value of their data (Tallon, 2013 p. 26).

It is noted that a mature digital state is a data-driven state that encompasses three major elements: nationwide data architecture, digital transformation of public administration and human resources. The shift to data-driven governance implies the creation of a unified state digital ecosystem with data and algorithms in the center building real-time data exchange networks to deliver improved public service. Such approach calls for rethinking of public administration principles unltimately ensuring the transition from an electronic state to a digital one to increase its efficiency. Thus, the effectiveness of the use of big data largely depends on the development of the digital environment and the ability to move data quickly and at low cost among a potentially unlimited number of participants in a government-regulated framework.

**Funding Information** Not applicable

## Declarations

The author has no relevant financial or non-financial interests to disclose.

**Conflicts of Interest** Not applicable

**Ethics Approval** Not applicable

**Informed Consent** Not applicable

## References

Akatkin Yu. M., Karpov, A. Y. M., Konyavskii, O. E., & Yasinovskaya, V. A. E. D (2017). Digital economy: conceptual architecture of a digital economic sector ecosystem. *Business Informatics*, *No. 4*(42), https://doi.org/10.17323/1998-0663.2017.4.17.28.

Barrett, J. (2022). (n.d.). *Up to 73% of Company Data Goes Unused for Analytics. Here's How to Put It to Work*. Inc. Retrieved May 29, from https://www.inc.com/jeff-barrett/misusing-data-could-be-costing-your-business-heres-how.html

Barton, D., & Court, D. (2012, October). *Making Advanced Analytics Work For You*. Harvard Business Review. Retrieved May 29, 2022, from https://hbr.org/2012/10/making-advanced-analytics-work-for-you

Bennett, S. (2017). What is information governance and how does it differ from data governance. *Governance Directions*, Sep. 2017. https://www.sibenco.com/wp-content/uploads/2017/09/Information_governance_data_governance_September_2017.pdf

Bernal, I., Sejersen, T., Hansen, R. M., & Nyasulu, A. M. (2021). Big data for economic statistics. *Stats Brief, UN ESCAP*, Issue No. 28, March 2021. https://repository.unescap.org/bitstream/handle/20.500.12870/3501/ESCAP-2021-PB-economic-statistics.pdf?sequence=1&isAllowed=y

*Big data is a job engine*. Bi-Survey. (n.d). Retrieved May 29, from https://bi-survey.com/big-data-investments

Big Data Association (2019). *Strategy for the Big Data market development by 2024*. Retrieved May 29, 2022, from https://rubda.ru/wp-content/uploads/2020/03/strategiya-bolshih-dannyh_srednyaya.pdf

Columbus, L. (2015, October 17). *Business Adopting Big Data, Cloud & Mobility Grow 53% Faster Than Peers*. Forbes. Retrieved May 21, 2022, from https://www.forbes.com/sites/louiscolumbus/2015/10/17/businesses-adopting-big-data-cloud-mobility-grow-53-faster-than-peers/?sh=5b76521f7d8a#7ff28d6e7d8a

Constatiou, I. D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57. https://doi.org/10.1057/jit.2014.17.

Davenport, T. H., & Dyché, J. (2013, May). *Big Data* in *Big Companies*. International Institute for Analytics. Retrieved May 23, 2022, from https://www.iqpc.com/media/7863/11710.pdf

Decree of the President of the Russian Federation No (2017). 203 "Strategy of the Information Society Development in the Russian Federation for 2017–2030" [Document]: May 9, 2017 – Moscow. Collection of the Legislative Acts of the Russian Federation. No. 20. Art. 2901

Dhamodharavadhani, S., Gowri, R., & Rathipriya, R. (2018, March). *Conference: First International Conference on Computer Vision,Networks and Informatics* Unlock Different V's of Big Data for Analytics. *International Journal of Computer Sciences and Engineering*, 06 (04), 183–190.

Dremel, C., Herterich, M. M., Wulf, J., Waizmann, J. C., & Brenner, W. (2017). How AUDI AG Established Big Data Analytics in its Digital Transformation. *MIS Quarterly Executive*, vol. 16 (2), art. 3, 81–100. https://aisel.aisnet.org/misqe/vol16/iss2/3

Federal Tax Service (2020, November 12). *Na obshchestvennom sovete pri FNS Rossii obsudili rabotu Sluzhby vo vremya pandemii [The Public Council of the Russian FTC discussed its work during the pandemic]* Retrieved May 21, 2022, from https://www.nalog.ru/rn77/news/activities_fts/10154705/

Goasduff, L. (2020, September 18). How Data and Analytics Leaders Can Master Governance. Gartner. Retrieved May 22, 2022, from https://www.gartner.com/smarterwithgartner/how-data-and-analytics-leaders-can-master-governance/

GOST R ISO/IEC 20546-2019. Information technology. Big data. Overview and vocabulary (2020). Federal Agency on Technical Regulating and Metrology. Moscow: Standartinform. https://api.bigdata-msu.ru/media/uploads/2020/05/06/1-025-20-20546-2019-end.pdf

GOST R 54524-2011/ISO/TS 8000-100:2009. Data quality. Part 100. Master data. Exchange of characteristic data. Overview (2012). Federal Agency on Technical Regulating and Metrology. Moscow: Standartinform. https://docs.cntd.ru/document/1200088822

GOST R 56214-2014/ISO/TS 8000-1:2011. Data quality. Part 1. Overview (2015). Federal Agency on Technical Regulating and Metrology. Moscow: Standartinform. https://docs.cntd.ru/document/1200114769

Gualtieri, M. (2016, January 21). *Hadoop Is Data's Darling For A Reason*. Forrester. Retrieved May 22, 2022, from https://go.forrester.com/blogs/hadoop-is-datas-darling-for-a-reason/

Heath, D. R. (2019). Prediction machines: the simple Economics of Artificial Intelligence. *Journal of Information Technology Case and Application Research*, *21*(3–4), 163–166. https://doi.org/10.1080/15228053.2019.1673511.

Hilb, M. (2019, May/June). *Unlocking the Board's Data-Value Challenge.Directorship*. NACD Directorship. Retrieved May 22, 2022, from https://www.nacdonline.org/insights/magazine/article.cfm?itemnumber=65499

Hunt, S. (2021, July 31). *Data Management Trends 2021*. Datamation. Retrieved May 24, 2022, from https://www.datamation.com/big-data/data-management-trends/

Ishwarappa, K., Anuradha, J. (2015). A brief introduction on Big Data 5Vs characteristics and Hadoop technology. *Procedia Computer Science*, *48*, 319–324. https://doi.org/10.1016/j.procs.2015.04.188.

Joshi, N. (2017, November 26). *Top 5 sources of big data*. Allerin. Retrieved May 24, 2022, from https://www.allerin.com/blog/top-5-sources-of-big-data

Judah, S. (2021, April 19). *Data and analytics now becoming a core business function*. SiliconAngle. Retrieved May 24, 2022, from https://siliconangle.com/2021/04/19/data-analytics-now-becoming-core-business-function/

Lynch, C. (2008). How do your data grow? *Nature*, *455*, 28–29. https://doi.org/10.1038/455028a.

Marr, B. (2020, September 21). *The Top 4 Artificial Intilligencce Trends For 2021*. Forbes. Retrieved May 24, 2022, from https://www.forbes.com/sites/bernardmarr/2020/09/21/the-4-top-artificial-intelligence-trends-for-2021/?sh=6433e4e31c2a

Mills, T. (2019, November 6). *Five Benefits of Big Data Analytics And How Companies Can Get Started*. Forbes. Retrieved May 24, 2022, from https://www.forbes.com/sites/forbestechcouncil/2019/11/06/five-benefits-of-big-data-analytics-and-how-companies-can-get-started/?sh=709e020b17e4

Milovidov, V. D. (2019). Faktory neopredelennosti mirovogo finansovogo rynka v usloviyakh tekhnologicheskoi revolyutsii [Global financial market uncertainty in the technological revolution]. Abstract of Doctor's degree dissertation. Moscow.

Mlitz, K. (2021, January 22). *Forecast revenue big data market worldwide 2011–2027*. Statista. Retrieved May 24, 2022, from https://www.statista.com/statistics/254266/global-big-data-market-forecast/

Moore, S. (2018a, June 19). *How to Create a Business Case for Data Quality Improvement* Gartner. Retrieved May 24, 2022, from https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement/

Moore, S. (2018b, January 18). *How to Stop Data Quality Undermining Your Business*. Gartner. Retrieved May 24, 2022, from https://www.gartner.com/smarterwithgartner/how-to-stop-data-quality-undermining-your-business/

OECD (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies. *OECD Publishing*. 2019, November 26. https://doi.org/10.1787/276aaca8-en

OECD (2019). The Path to Becoming a Data-Driven Public Sector. *OECD Digital Government Studies*, *OECD Publishing*. 2019, November 28. https://doi.org/10.1787/059814a7-en

OECD (2019). Strengthening Digital Government. *OECD Going Digital Policy Note, OECD*. 2019, March 7. Retrieved May 24, 2022, from https://www.oecd.org/going-digital/strengthening-digital-government.pdf

OECD (2020). The OECD Digital Government Policy Framework: Six dimensions of a Digital Government. *OECD Public Governance Policy Papers*, *OECD Publishing*, 2020, October 7, No. 02. https://doi.org/10.1787/f64fed2a-en

Petrov, M., Burov, V., Shklyaruk, M., & Sharov, A. (2018, April). *The State as a Platform: People and Technology*. Center for Strategic Research. Retrieved May 24, 2022, from (https://www.csr.ru/upload/iblock/313/3132b2de9ccef0db1eecd56071b98f5f.pdf)

Pharmaceutical technology *Big data and AI are likely to lead emerging technology investments: Survey*. Retrieved May 25, 2022, from https://www.pharmaceutical-technology.com/surveys/big-data-and-ai-are-likely-to-lead-emerging-technology-investments-survey/

Precisely (2021). *Data Integrity Trends: Chief Data Officer Perspectives in 2021* Retrieved May 25, 2022, from https://www.precisely.com/resource-center/analystreports/data-integrity-trends?utm_source=Referral&utm_medium=Press-Release

Rialti, R., Ciappei, C., Zollo, L., & Boccardi, A. (2016). June 9–10). *Conference: XXVIII Sinergie Annual Conference* Big data oriented business models: the 7vs of value creation. Vol. Management in a Digital World. Decisions, Production, Communication, pp. 101–103. University of Udine. Italy. https://www.researchgate.net/publication/310479231_Big_data_oriented_business_models_the_7vs_of_value_creation

SAP (2021). *Real-Time, Data-Driven Government. Develop Forward-Thinking, Citizen-Centric Programs*. The SAP Institute For Digital Government. Retrieved May 25, 2022, from https://www.sap.com/documents/2018/05/dce5d788-057d-0010-87a3-c30de2ffd8ff.html

SAS. *Big data: what is it and why it matters* Retrieved May 25 (2022). from https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Sicular, S. (2013, April 2). *Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s*. Gartner. Retrieved May 25, 2022, from https://blogs.gartner.com/svetlana-sicular/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/

Statista Research Department (2021, Seprember 8). *Volume of data/ information created, captured, copied, and consumed worldwide from 2010 to 2025*. Statista. Retrieved May 24, 2022, from https://www.statista.com/statistics/871513/worldwide-data-created/

Tallon, P. P. (2013). Corporate governance of Big Data: perspectives on Value, Risk, and cost. *Computer*, *46*(6), 32–38. https://doi.org/10.1109/mc.2013.155.

The Economist *The world's most valuable resource is no longer oil, but data: Regulating the Internet giants*. Retrieved May 24, 2022, from https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

The Russian Presidential Academy of National Economy and Public Administration (2019). *Iskusstvennyi intellekt. K vyboru strategii [Artificial Intilligence. Towards a choice of strategy]*. The education center for digital transformation teams and CDTOs. Retrieved May 24, 2022, from https://www.ranepa.ru/images/News/2019-05/30-05-2019-AI-report-2019.pdf

The Russian Presidential Academy of National Economy and Public Administration (2019). *Gosudarstvo kak platforma: Lyudi i tekhnologii [The state as a platform. People and technology]* The education center for digital transformation teams and CDTOs. Retrieved May 24, 2022, from https://files.data-economy.ru/Docs/GovPlatform.pdf

Ubaldi, B., Le Fevre, E. M., Petrucci, E., Marchionni, P., Biancalana, C., Hiltunen, N., Intravaia, D. M., & Yang, C. (2019). State of the art in the use of emerging technologies in the public sector. *OECD Working Papers on Public Governance*, *OECD Publishing*. No. 31. https://doi.org/10.1787/932780bc-en

VentureBeat *Precisely: 82% of data executives cite data quality as a barrier*. Retrieved May 24, 2022, from https://venturebeat.com/2021/06/25/precisely-82-of-data-executives-cite-data-quality-as-a-barrier/

World Bank (2021). World Development Report 2021: Data for Better Lives. https://doi.org/10.1596/978-1-4648-1600-0

Yukhno, A. S. (2019). Prospects for the use of artificial intelligence in corporate governance. *Insurance law*, *no. 4*(85), 26–32.

Yukhno, A. S. (2021). Corporate governance of information and communication technologies in the digital era. *Bulletin of the Institute of Economics of the Russian Academy of Sciences*, (6), 127–145. https://doi.org/10.52180/2073-6487_2021_6_127_145.