

## The Notion of Big Data and its Legal Aspects

Sarvar Tojiboyev

Tashkent State University of Law, Uzbekistan



DOI : <https://doi.org/10.61796/ijblps.v2i4.302>



### Sections Info

#### Article history:

Submitted: February 19, 2025

Final Revised: February 26, 2025

Accepted: March 05, 2025

Published: March 12, 2025

#### Keywords:

Big data

Legal regulation

Notion of big data

Technology

Information

Complex institute

Legal phenomenon

### ABSTRACT

**Objective:** The objective of this study is to highlight the rapid growth of data generated by society through the widespread use of the Internet and the need for new technologies to process large volumes of information. The paper examines how commercial organizations, especially giant companies with access to advanced technologies, have an advantage in data processing, while smaller organizations struggle to compete. **Method:** The paper reviews the IDC report, "The Data Age 2025," which outlines the trends in the information technology market until 2025. It emphasizes the projections regarding the growth in data volume, particularly the 60% of information expected to be produced by commercial organizations. Additionally, the paper discusses the application of Big Data technology, specifically in Amazon's automated system that uses customer data to personalize product offerings. **Result:** The study finds that by 2025, the total volume of data produced will reach 175 zettabytes, with a substantial portion of it generated by commercial entities. The increasing reliance on Big Data and automated systems will give companies with access to large-scale technologies a significant advantage. Organizations without access to these technologies may face challenges in competing due to their inability to process large amounts of data effectively. **Novelty:** The novelty of this paper lies in the exploration of the pressing need for advanced data processing technologies in today's data-driven society. It brings attention to the role of commercial companies, like Amazon, in utilizing Big Data to create personalized customer experiences, offering a glimpse into how data technologies are shaping future market dynamics and organizational strategies.

## INTRODUCTION

Initially, it is appropriate to get closely acquainted with this technology. According to scientific literature, the concept of "big data" was first introduced to science by Clifford Lynch. In 2008, in the British scientific journal Nature, he discussed how technologies that open up opportunities to work with large amounts of data might affect the future of science, indicating the emergence of a new paradigm in science [1]. However, the concept of Big Data was also recognized by scientists such as John R. Mashey, Michael Cox, and David Ellsworth in scientific developments before Clifford Lynch [2]. Big Data has been growing in popularity each year, attracting the attention of many scientists. By 2011, besides Amazon, Microsoft and IBM also began using Big Data in their projects, which further increased interest in this new phenomenon. By 2013, many foreign higher education institutions began introducing independent courses dedicated to Big Data. Consequently, this reinforced the need to analyze this technology-based new concept not only from a technical perspective but also from a legal standpoint [3]. Today, many legal scholars view the concept of Big Data as a large database consisting of various types of data or information. One of the main reasons for this preliminary conclusion is the definition given by Clifford Lynch. The author, in turn, associated this phenomenon with

the sharp increase in global information volume and data diversity [4]. Therefore, like many others, legal scholars also have a conventional view of Big Data, evaluating it as a collection of numerous and diverse data sets. In our opinion, such considerations and views are not sufficiently globalized and do not have their proper place in science. The rapidly growing need for Big Data today undoubtedly indicates the necessity not only for its technical regulation but also for creating legal regulatory aspects. However, it is worth noting that in recent years, jurisprudence has been unable to keep pace with the latest technological innovations. In particular, along with Big Data, global concepts such as IoT (Internet of Things), startup, Text and Data Mining (TDM), non-relational (noSQL) database have not been subjected to a unified legal regime [5].

## RESEARCH METHOD

Considering the complexity of changing the legal system, it is appropriate to revitalize the legal regulation of these areas. First and foremost, taking into account the dynamic and constantly changing nature of the innovation sector, it is necessary to create a specific legal construction for the new phenomenon. This legal construction should not only reflect the basic regulatory procedures for Big Data but also serve as a foundation for developing future legislation in this field. Therefore, it is first necessary to identify the main characteristics and legal nature of the Big Data concept. For example, V. Vaypan emphasizes the need to form fundamental legal concepts and institutions when studying the digital economy by periods [6]. There have been several attempts to develop definitions for the Big Data concept in literature and the legal field, but these approaches are proving inadequate for the current situation. In Internet networks and literature, there are various considerations about Big Data, which are not interpreted based on a single approach. Particularly, according to D. Laney and J. Manyika's views, "large volumes of data" are large quantities of rapidly changing and/or diverse information sources that require modern processing methods to improve decision-making, discover new knowledge, and ensure process optimization [7]. According to B. Jacques, Big Data is considered to be data that, due to its volume, forces humanity to move beyond traditionally widespread and tested methods of that time [8]. Field scholar M. Loukides suggests that Big Data technologies can help when the volume of data itself becomes part of the problem and traditional methods of data processing become ineffective [9]. In our view, the above scholars mainly focused on discussing the conditions in this process, its methods, and their relationships with traditional data collection methods, rather than directly defining the essence of Big Data. M. Rozhkova evaluates Big Data as a constant flow of large volumes of data from various sources. Here, the author refers to information exceeding 100 GB per day [10]. In our opinion, this approach is considered subjective, as experts mainly emphasize the volume of data. In addition to the above, various scientific and international communities have attempted to clarify views on Big Data, primarily focusing on its technological aspects. Specifically, according to NESSI community reports, Big Data is a concept that includes methods of collecting, processing, analyzing, and visualizing large volumes of data within acceptable time frames that cannot be

processed using standard information technologies. The platforms, tools, and software used for this purpose are collectively referred to as Big Data technologies [11]. According to the Gartner network, specialized in ICT, Big Data consists of information resources with large volume, high velocity, and/or variety, requiring effective and innovative data processing methods. These methods expand the possibilities for deeper understanding, correct decision-making, and process automation [12]. According to the Techopedia community, large volumes of data represent a process used when traditional data "mining" and processing methods cannot determine the essence and meaning of the main data. Current relational database systems cannot process unstructured, time-dependent, or simply very large volumes of data. This type of data requires a different processing approach called big data, which uses massive parallelism in easily accessible hardware [13]. In the European Parliament's 2021 ICT reports, Big Data refers to data sets of such volume and complexity that they are difficult to process with traditional ICT applications [14]. Additionally, McKinsey Global Institute's 2011 report "Big data: The next frontier for innovation, competition, and productivity" defined Big Data as a collection of data that increases the volume of collecting, storing, managing, and analyzing traditional databases through software tools [15].

## RESULTS AND DISCUSSION

### *Results*

Based on the above concepts, it can be said that Big Data is a term representing such large and complex data sets that processing them using existing database management tools or traditional data processing programs becomes difficult. Another group of representatives evaluate Big Data as a collection of technologies [16] that perform several functions. Here, the operations related to technology are significant as they include processes of working with large volumes of both structured and unstructured data in various fields, distinct from traditional data.

The third group of representatives discusses the complex nature of Big Data, viewing it as both a "technology and database" for data processing. The fourth group envisions Big Data as directly "related to intellectual property law (IP law)" [17] and evaluates its existence on this basis.

Taking into account the views and considerations of the above groups, it is appropriate to conditionally divide the perspectives on Big Data into three groups:

1. As a technology;
2. As data;
3. As a complex institute combining data and technologies;
4. As a legal phenomenon.

The first group representatives recognize Big Data as "technology". However, global experience [18] emphasizes that in large digital data research analysis, it is the data, not the technologies that are primary. In this direction, besides certain rules established in national legislation, foreign state experiences and field specialists have also presented argumentative views and evidence. Undoubtedly, it should be noted that the

national legislation has not yet developed a definition for the Big Data concept, and until now, in legal practice, it has been accepted as a simple collection of data

### *Discussion*

In the explanatory dictionary of Information Technology terms [19], we can see that there is no separate definition of large digital data. In general terms, this dictionary explains that large volumes of data are related to the abundance of data scale. Article 3 of the Law of the Republic of Uzbekistan on Informatization defines the concept of “information technology” as all methods, devices, techniques, and processes used for collecting, storing, searching, processing, and distributing information [20]. In our opinion, this rule in the law indicates that information technology is a large and general concept, and Big Data is not specifically technology.

Moreover, characteristics typical of any technology do not match the content of Big Data, including:

1. Existence of algorithmic processes and ability to manage them;
2. Presence of internal operational structure;
3. Use of specific methods and mechanisms in technological processes;
4. Existence of purpose and focus on achieving it.

Thirdly, in the world of information technology, there are important terms reflecting the processes of data processing, analysis, and systematization, without which it is difficult to imagine Big Data operations. In particular, according to

D. Gervais, Big Data is developed and created in two ways [21]: initially, creating a Big Data database (corpus), which requires processes of collecting data from various open sources. Secondly, the database collected for Big Data is analyzed through the “Text and Data Mining (TDM)” process [22]. In the TDM process, artificial intelligence algorithms are certainly used. According to C. Kozyrkov, this situation allows the “machine” to conduct analyses from the database, which is why sometimes the term “Machine learning” (a field of artificial intelligence aimed at developing independent learning programs) is also used as a synonym for artificial intelligence in the publishing field [22]. Indeed, as the machine analyzes the Big Data database, it learns it better and presents more perfect results. Additionally, the term “Deep learning” (aimed at mastering complex independent programs) also plays a key role in the Big Data process.

### **CONCLUSION**

**Fundamental Finding:** The fundamental finding of this study is the proposed definition of Big Data, which is presented as a collection of data that significantly enhances the processes of data collection, storage, management, and analysis through modern software tools and innovative methods. The research underscores that Big Data is a constantly evolving and large volume of various types of data, which brings about a shift in how we perceive and understand the “data world.” **Implication:** The implication of this finding is that the rapidly growing volume and complexity of Big Data require new approaches and technologies for effective processing and analysis. This presents

both opportunities and challenges for organizations and researchers, as they must adopt innovative methods to handle the ever-increasing amounts of diverse data. The new information reality facilitated by Big Data will likely reshape industries, decision-making processes, and societal norms. **Limitation:** A limitation of this study is that while the definition of Big Data is proposed, it does not delve deeply into the specific technical challenges organizations face when adopting Big Data technologies. Additionally, the study does not explore the potential ethical implications and privacy concerns related to the use of large-scale data collection and analysis. **Future Research:** Future research should focus on exploring the practical applications of Big Data in various industries, with particular attention to overcoming the technical barriers organizations face in data management and analysis. Additionally, further investigation into the ethical concerns surrounding data privacy, security, and regulation in the context of Big Data is needed. Exploring the potential impact of emerging technologies such as artificial intelligence and machine learning in Big Data analysis would also be valuable for understanding the future directions of this field.

## REFERENCES

- [1] L. Chernyak, "Big Data - new theory and practice," Open Systems, no. 10, 2011. M. S. Korneev, "History of the concept of 'Big Data': dictionaries, scientific and business periodicals," RSUH Bulletin, 2018. N. I. Tsarkova and A. S. Smolyanov, "Big Data. Development, analysis and technology," Actual Problems of Humanities and Natural Sciences, no. 7-1, 2016.
- [2] J. R. Mashey, "Big Data and the Next Wave of Infra Stress," 1998. [Online]. Available: <https://www.usenix.org/conference/1999-usenix-annual-technical-conference/big-data-and-next-wave-infrastress-problems>. [Accessed: 03.09.2018]. M. Cox and D. Ellsworth, "Application-controlled demand paging for out-of-core visualization," Proceedings of the IEEE 8th Conference on Visualization, Oct. 1997. [Online]. Available: <https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf>. [Accessed: 03.09.2018].
- [3] P. Lyman and H. R. Varian, "How Much Information," 2003.
- [4] D. Reinsel, J. Ghantz, and J. Rhyding, "The Digitization of the World. From Edge to Core," The Data Age 2025. [Online]. Available: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagatedataage-whitepaper>.
- [5] "Internet publication," Nature, [Online]. Available: <https://www.nature.com/articles/455001a>.
- [6] I. Savelyev, "Problems of applying personal data legislation in the era of 'Big Data'," Law, Journal of Higher School of Economics, no. 1, 2015, pp. 43-66.
- [7] V. A. Vaypan, "Fundamentals of Legal Regulation of Digital Economy," Law and Economics, no. 11, 2017, p. 8.
- [8] J. B., R. D., C. R., and A. H. B., "Big Data: The next frontier for Innovation, Competition, and Productivity," McKinsey, 2011. [Online]. Available: [www.McKinsey.com](http://www.McKinsey.com). [Accessed: 23.02.14].
- [9] M. Loukides, *The Evolution of Data Products*, O'Reilly Media, Inc., 2011.

- [10] M. Rozhkova, "What is Big Data?" [Online]. Available: [https://zakon.ru/blog/2019/4/22/chto\\_takoe\\_bolshie\\_dannye\\_big\\_data](https://zakon.ru/blog/2019/4/22/chto_takoe_bolshie_dannye_big_data).
- [11] NESSI White Paper, "Big Data A New World of Opportunities," Dec. 2012.
- [12] "Internet source," Gartner, [Online]. Available: <https://www.gartner.com/en/information-technology/glossary/big-data>.
- [13] "Internet source," Techopedia, [Online]. Available: <https://www.techopedia.com/definition/27745/big-data>.
- [14] "Internet source," European Parliament, [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20210211STO97614/big-data-definition-benefits-challenges-infographics>.
- [15] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute, 2011. [Online]. Available: <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-nextfrontier-for-innovation>.
- [16] S. Protasov, "What is Big Data?" PostNauka. [Online]. Available: <https://postnauka.ru/faq/46974>.
- [17] D. Gervais, "Exploring the Interfaces Between Big Data and Intellectual Property Law," JIPITEC, vol. 10, 2019, p. 22.
- [18] C. Snijders, U. Matzat, and U.-D. Reips, "'Big Data': Big gaps of knowledge in the field of Internet," International Journal of Internet Science, vol. 7, no. 1, pp. 1-5, 2012. [Online]. Available: [http://www.ijis.net/ijis7\\_1/ijis7\\_1\\_editorial.pdf](http://www.ijis.net/ijis7_1/ijis7_1_editorial.pdf).
- [19] National Database of Legislative Information of the Supreme Assembly of the Republic of Uzbekistan, Document No. 03/21/683/0375, adopted April 21, 2021.
- [20] D. Gervais, "Exploring the Interfaces Between Big Data and Intellectual Property Law," JIPITEC, vol. 10, 2019, p. 23.
- [21] M. Lilla Montagnani, "Il text and data mining e il diritto d'autore," AIDA, vol. 26, 2017, p. 376.
- [22] C. Kozyrkov, "Are you using the term 'AI' incorrectly?" Hackernoon, 26 May 2018. [Online]. Available: <https://hackernoon.com/are-you-using-the-term-ai-incorrectly-911ac23ab4f5>.

---

**\*Sarvar Tojiboyev (Corresponding Author)**

Tashkent State University of Law, Uzbekistan

Email: [szafarovic@gmail.com](mailto:szafarovic@gmail.com)

---