The Features of Legal Regulation of Big Data in the Context of the Development of the Digital Earth

Sona De Apro[*][1]

[1]Lomonosov Moscow State University, Moscow, Russia, S.Deapro@yandex.ru

Abstract

When creating a Digital Earth, huge amounts of data are generated through remote sensing and various sensors to observe our planet. Big Data has become of particular value in the context of the development of the Digital Earth. At the same time, every day there is more and more information, which creates serious problems for their collection, storage, processing, ensuring their security and legal regulation. The advent of cloud technologies has made it possible to provide computing support for collecting, storing and processing Big Data, while the heterogeneity of the amount of data collected is growing and it becomes difficult to determine their legal regime, jurisdiction and protection in general. The report analyzes the sources of Big Data used for the development of the Digital Earth, the peculiarities of their collection, storage and processing through cloud technologies, analyzes the prospects and risks of the development of the Digital Earth in terms of the security of Big Data and their legal regulation.

Keywords

Digital Earth, Big Data, Cloud Technologies, Legal Regulation, Information Security, IT.

Introduction

The rapid development of the digital economy every day leads to an increase in the amount of data, its heterogeneity and unstructuredness. In this connection, the term Big Data appeared in popular science terminology, which is attributed to almost any data collected from a variety of sources. In the Strategy for the development of the information society of the Russian Federation for 2017-2030[1], it is noted that the processing of Big Data is one of the most important issues in the context of the development of information technologies. To collect, process, store such data, cloud technologies are used, which are also among the priorities in terms of the development of the digital economy and the Digital Earth.

It should be noted that the term Big Data, as well as its definition, is not legally enshrined either in Russian legislation or in other foreign laws. In the Russian law on personal data, there are concepts such as personal data, biometric personal data. The GDPR law includes similar concepts, as well as genetic data, data related to health. In the United States, there is no federal regulation of the term Big Data. Despite the status of the most advanced states in the world in terms of digital technologies, the United States does not have a single centralized approach to regulating the protection of personal data and, moreover, Big Data. The rules governing the legal procedure for the

---

[*]   Corresponding author
[1]   Decree of the President of the Russian Federation of 05.09.2017 № 203 «On the Strategy for the Development of the Information Society in the Russian Federation for 2017 – 2030»

use of such data are scattered across many acts, for instance:

- The Fair Credit Reporting Act (15 U.S.C. §1681);

- The Health Insurance Portability and Accountability Act (HIPAA) (42 U.S.C. §1301);

- The Financial Services Modernization Act (15 U.S.C. §§6801-6827);

- The Electronic Communications Privacy Act (18 U.S.C. §2510);

- The Computer Fraud and Abuse Act (18 U.S.C. §1030).

The difficulty of defining the term Big Data is understandable. After all, data collection is carried out from a variety of sources. These include remote sensing of the Earth, geotechnology, face recognition technology, the IoT, mobile phones and other electronic devices, the Internet, distributed registry technologies. Entertainment content and video surveillance have long been important sources of growth in the global data space. As we move digital, there is a growing share of data from IoT devices, metadata (critical for analytics, contextualization, and artificial intelligence), and operational data. Four industries have the greatest potential for data generation: healthcare, industry, financial sector, and media business. In addition, Big Data is structured and unstructured information from a large number of different, including scattered or loosely coupled, sources in volumes that cannot be processed manually in a reasonable time. According to the representatives of the expert community, structured data is understood as ordered information contained in databases, information systems, etc. And unstructured data is information contained in audio and video recordings, text included in graphic images, that is, data in formats making it difficult to find specific information in them.

Discussion

According to a late 2018 study by IDC, the global data sphere will more than fivefold by 2025, to 175 ZB, up from 33 ZB in 2018[2]. At the same time, the volume of the global data sphere is determined by the volume of new data created and reproduced per year. Storage is also expected to undergo significant changes. The share of data stored in public clouds is increasing, and by 2025 this share will reach 45-50%[3]. This leads to the complication of centralized collection, storage and manual processing of data. Moreover, cloud technologies are directly related to the Internet, and here comes the jurisdictional aspect of storing data about citizens of different states and the issue of

---

[2] [Electronic resource] URL: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

[3] [Electronic resource] URL: http://www.garant.ru/article/1263975/

responsibility for the confidentiality and security of this data. Cloud technologies are used to store certain data in absolutely all areas, and as noted earlier, the development of cloud technologies is a priority among the tasks in the development of the digital economy and the creation of the most accurate picture of the Digital Earth.

The increasing availability of Big Earth data has provided unprecedented opportunities to understand the Earth in the Digital Earth context. In recent research, Big Data have been characterized by 5 Vs (volume, variety, velocity, veracity, and value).[4] «Data collection strategies, data storage facilities, data analysis methods, and data access services facilitate the transformation from the other 9Vs to the 10th V of value. With the continuing increases in the volume and complexity of data, there are challenges in the life cycle of data management, including data storage, data query, data analysis, data sharing, and many other aspects. Managing Big Data requires an extensible, interoperable and scalable architecture that supports data storage and analysis. Fortunately, recent years have witnessed the evolution of cloud computing, which brings potential solutions to support the life cycle of big data management».[5]

Together with the term «cloud technologies» in regulatory legal acts, there are such terms as: cloud computing, cloud services, cloud, distributed information systems, service-oriented information systems, service-oriented architecture, software as a service, software rental, applied information systems of the «SaaS» model, cloud computing, consulting in the field of information services and a number of others. Undoubtedly, cloud technologies are gaining more and more popularity and more and more advanced formats for collecting and storing data in the cloud computing environment appear. The essence of cloud technologies boils down to the fact that instead of acquiring and using their own computing equipment, they directly use the computing resources of a company that provides services based on cloud technologies. The advent of cloud technologies has made it possible to provide computing support for collecting, storing and processing Big Data, while the heterogeneity of the amount of data collected is growing and it becomes difficult to determine their legal regime, jurisdiction and protection in general. In this connection, there is a risk of hacking, theft and abuse of the information received. Today there is no single organization that would centrally control and conduct regular audits of available and accumulating data from different sources. Some subordinate bodies have similar powers. For

---

[4]   Marr B (2015) *Big data: using SMART big data, analytics and metrics to make better decisions and improve performance*. John Wiley & Sons, Hoboken, NJ

[5]   Yun Li, Manzhu Yu, Mengchao Xu, Jingchao Yang, Dexuan Sha, Qian Liu and Chaowei Yang (2019) *Chapter 19: Big Data and Cloud Computing* In: Annoni A, Goodchild M, Guo H. (eds), Manual of Digital Earth, Springer, pp. 325-357. DOI:10.1007/978-981-32-9915-3_23

example, in November 2020, the Russian Ministry of Internal Affairs announced that it plans to create a centralized bank of biometric data of Russians and foreign citizens by 2023. According to the Ministry of Internal Affairs, the technology will allow identification of a person, as well as unidentified bodies by fingerprint information, facial images and genomic information. The development will interact with the combined search federal system for genetic identification. However, due to the fact that the foundation of any technology is foreign development, then here too one has to face restrictions in the field of full access to the necessary information. That is why in Russia it is planned to switch companies to domestic software development in order to avoid the dependence of Russian companies on foreign software and electronics.

Conclusions

The main problems of legal regulation of Big Data are lack of classification and identification of Big Data and lack of understanding of the limit and boundaries of data collection and establishment of its legal regime. In addition, the concept of Big Data is not defined and legislatively fixed. The problems of cloud technologies are that the subjects of cloud technologies are not defined, the security and privacy of Big Data stored in the cloud is vulnerable and the boundaries of jurisdictions are not defined. Thus, legal regulation of Big Data and cloud technologies must take into account the jurisdictional aspect and data centralization. The prospects for the development of cloud technologies and Big Data are clear. It is obvious that cloud technologies will continue to be used everywhere, new capacities will be developed, centers for collecting, processing and storing information will be created. New software will evolve for high-performance and reliable storage systems, leading to the rapid development of the Digital Earth and the ability to analyze the most complex data and extract conclusions in real time.

It is crucial to identify as accurately as possible both the data and the sources that produce and process it to build an organized Big Data environment. In this case, more precise control and data audit will be possible. To develop the Digital Earth and ensure the confidentiality and security of databases and data themselves, it is necessary to identify and classify data and their sources, introduce into legislation the concepts of Big Data and its definitions, designate the legal regime, conduct regular audits and implement control mechanisms for accessing Big Data. There is the need to transfer of databases and cloud technologies, as well as from subsequent collection, storage and processing to domestic developments, as well. For the objectivity of the collection, processing,

storage, audit and control of databases without abuse of one's own powers, there is an urgent need to create an independent organization to protect the confidentiality of Big Data in the era of development of information technology, the digital economy and the Digital Earth.

References

Yun Li, Manzhu Yu, Mengchao Xu, Jingchao Yang, Dexuan Sha, Qian Liu and Chaowei Yang (2019) *Chapter 19: Big Data and Cloud Computing* In: Annoni A, Goodchild M, Guo H. (eds), Manual of Digital Earth, Springer, pp. 325-357. DOI:10.1007/978-981-32-9915-3_23

Marr B (2015) *Big data: using SMART big data, analytics and metrics to make better decisions and improve performance*. John Wiley & Sons, Hoboken, NJ

Savelyev AI (2007) *Features of legal regulation of the civil status of Internet access providers // Actual problems of Russian law.* No 1.

Decree of the President of the Russian Federation of 05.09.2017 № 203 «On the Strategy for the Development of the Information Society in the Russian Federation for 2017 – 2030»

The Digitization of the World (November 2018). URL: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf Accessed 30 Dec 2020

Data Management in the Transition to the Digital Economy (2019, March 19) URL: http://www.garant.ru/article/1263975/ Accessed 30 Dec 2020