

Legal and environmental problems of personal data protection under commercialization of Big Data

A. A. Vasiliev*, Yu. V. Pechatnova

Altai State University, Barnaul, Russian Federation

*Corresponding author E-mail: anton.vasiliev@mail.ru

Received: 15.09.2020. Accepted: 20.10.2020

The article deals with the legal-environmental problems of personal data protection related to the use of artificial intelligence-based programs of information collection, processing and storage. The normative and legal acts on personal data protection are presented, attention is paid to the international experience of legal regulation of this field. The scope of the changes that have occurred in the field of personal data protection due to the appearance and rapid increase of the "Big Data". The materials of the jurisprudence on the commercialization of the "Big Data" and, as a consequence, the violation of the existing legal regime of personal data protection are analyzed. In conclusion, a legal assessment of the risks of deanonymization of personal data and possible options to address the problem of implementation of personal data legislation in an era of rapid development of Big Data are proposed.

Keywords: Personal data; Environmental; "Big Data"; Information protection; Artificial intelligence

Introduction

The newest technologies open huge opportunities for creating advanced methods of collecting, processing and storing information. These programs, called "Big Data" in the business environment, can generate results that were not covered by the developer's intention, but were made possible by the introduction of the ability to analyze and compare an unlimited amount of data.

The formation of the term "Big Data" was facilitated by its use by Clifford Lynch to describe the phenomenon of rapid growth of the world's volumes of information (Rozhkova, 2019).

In general, "Big Data" has the following features:

- 1) Are formed using the latest database management systems, therefore, Big Data is not so much information itself as a symbiosis of information and its processing technologies;
- 2) Represent a new form of knowledge production based on revealing connections and regularities between data, establishing hidden trends and making forecasts;
- 3) Are a dynamic continuous process of data processing on a global scale (according to statistics, by 2020 the accumulated volume of data should increase to about 44 zettabytes or 44 trillion gigabytes (in 2015 it was 12 zettabytes) (Volume of data, 2019);
- 4) Arise spontaneously as a by-product of continuous information generation; in other words, they are secondary data that were originally intended for some purposes but subsequently continue to be generated, transformed and reused for other purposes;
- 5) In contrast to traditional databases, the main measurement of which is the number of information units, to characterize the "Big Data" (Big Data) used several measurements - 3Vs: volume (physical volume); velocity (growth rate); variety (variety of data).

Methodology

The rapid growth rate of "Big Data", high relevance of the issue of developing effective legislation in the use of artificial intelligence predetermined the formation of various proposals for legislative regulation of this issue at both the national and international levels.

In the Russian Federation, within the framework of the Digital Economy Program aimed at developing the country's innovation potential, Big Data is taking the leading position as a key digital technology. At the same time, Russian legislation does not yet introduce the term "Big Data" (Big Data), nor does it regulate the relationship between "Big Data" and personal data (Sosnin, 2019).

At the interstate level, many proposals on the regulation of information protection are not in favor of strengthening measures for personal data protection. Thus, the European Commission's communication "Artificial Intelligence for Europe" outlines the position of the European Union, which shows a tendency to narrow the human right to own and manage information about oneself: "...public policy should encourage greater availability of private data..." (Uroshleva, 2018).

This promotion of Big Data development on an interstate scale is explained by the fact that information flows have a significant impact on all sectors of the world economy. As a result, the Big Data market is gradually becoming an important catalyst for economic growth.

Results

There are many examples of successful use of Big Data. For example, Spaceknow turns to the use of images from space for oil price forecasting: The analysis algorithms are able to determine the level of filling of the largest oil storage facilities depending on the angle of shadow fall (Rozhkova, 2019).

In the sphere of banking activity there is an increase in the activity of the industry of information brokers, which through the Big Data technologies accumulate and analyze information about potential borrowers in the interests of lenders (Saveliev, 2015).

In addition, Big Data has a significant commercial value for insurance companies in order to analyze the probability of occurrence of an insured event, to assess the insurance risks and to determine the adequate amount of insurance premiums (Saveliev, 2015).

Thus, the global turnover of Big Data is the most important factor and necessary condition for the development of digital economy. At the same time, significant changes in the field of information turnover affect the relations between the exchange of personal data and protection of private life in such a way that the latter require taking additional measures for their protection.

According to the Federal Law of 27.07.2006 No. 152-FZ "On personal data", personal data are any information relating to a directly or indirectly defined or identifiable natural person (Federal Law, 2006). Based on this definition, we can conclude that personal data are part of the "Big Data".

At present, personal data is limited in its negotiability due to the public legal regime for its protection. At the same time, the development of digital technologies demonstrates a long-lasting tendency to decrease the protection capacity of the institution of personal data protection, which aggravates the conflict between the advantages of modern technologies and the right to privacy.

The use of personal data in most cases implies their intended use. However, given the fact that "Big Data" by its characteristic is a digital reusable asset, the legalization of their intended use hinders business development.

Therefore, because Big Data is gradually being commercialized, personal data periodically becomes a currency for paying for Internet services, an asset of companies, a marketing maneuver, a commercial product provided by information brokers to interested companies.

Some experts point out that the practice of using "Big Data", including personal data, is inevitable, so the possibility of legal introduction of personal data into civil circulation should be envisaged. Among the arguments in favor of this measure are the following: firstly, to prevent the appearance of a shadow market of personal data and to harmonize the turnover of personal data with the legislation; secondly, to remove the barriers to building a digital economy; thirdly, the narrowing of personal data content and changing the way of their use does not mean the violation of privacy (Uroshleva, 2018).

As an alternative measure, it is proposed to transfer to the civil circulation of user data, while personal data should be left infringing. However, the introduction of this measure is complicated by the imperfection of the conceptual framework: the ratio of categories of information, "Big Data", user data, personal data remains unclear (Uroshleva, 2018).

As such, information is an object of civil rights.

Personal data is a type of information, which cannot act as an object of civil law transactions. At the same time, personal data are not completely excluded from civil transactions, as evidenced by the possibility to use personal data with the consent of the personal data subject.

User data are a set of information about natural persons and their behavior that does not contain personal data, which does not allow to identify a particular natural person without using additional information or additional processing. The totality of user data is also often defined as "digital personality". Based on the definition, user data can be subject to civil circulation.

At the same time, it is not possible to distinguish between user data and personal data, to introduce some of them and to strengthen protection measures against the circulation of others. The key risk is the possibility of de-anonymization of data by artificial intelligence, which has the ability, when analyzing several impersonalized user databases together, to establish the personal data of a specific person (Voynikanis et al., 2018).

Thus, the two main factors: the increase in productivity and availability of artificial intelligence computing power, as well as a huge amount of personal information available on the Internet - determine the inevitability of civil circulation of user data and, consequently, the technical possibility of de-anonymizing even carefully impersonalized personal data.

Several practical examples show that, to date, there is no uniform jurisprudence on the issue of the turnover of user data and the protection of personal data.

Firstly, the company "Vkontakte" in January 2017 filed a lawsuit for the protection of rights to the database of social network users to the company "Double Data" and the National Bureau of Credit Histories. According to the claimant, the defendants extracted the data of social network users and processed them for commercial purposes. Objecting to the satisfaction of the claim, the company "Double Data" stated that "Vkontakte" has no exclusive rights to the created database. Because for the social network this database is not the main but a by-product, it is possible to limit the use of publicly available data only by federal law.

The court of first instance dismissed the claims of Vkontakte. The court of appeal noted that "Vkontakte" has exclusive rights to the database, the organization and maintenance of which the company bears the costs. Therefore, the decision of the court of first instance was overturned and a new decision was made in the case, which fully satisfied the claims of Vkontakte.

The court of cassation overturned the decision of the court of appeal, which prohibited "Double Data" from using open data "Vkontakte" for commercial purposes. The court, considering the dispute, did not consider proven the fact that "Double Data" extracted user data from the database collected in the social network, because the technology works on the principle of search engines, and open data are accumulated from various resources. In this regard, the case was sent for new consideration (Brezgulevskaya, 2018).

Secondly, the dispute between "HeadHunter" and "Staphori" Ltd. has become similar, long and ambiguous. According to the plaintiff's arguments, the service of LTD "Stafori" took commercial advantage of the CV database of the resource hh.ru and provided access to the data to third parties without the consent of the personnel portal (Brezgulevskaya, 2018).

Conclusions

Big Data technology marked the moment when the notion of post-industrial (information) society acquired its full meaning. Information has acquired the status of a valuable asset. On the one hand, Big Data technologies have enormous commercial value for building new business models. On the other hand, Big Data has significant potential to invade the privacy of citizens.

The current legislation presupposes that an individual has the possibility to dispose of personal data independently, while the processing of personal information is carried out only with the consent of the personal data subject.

At the same time, "Big Data" is incompatible with the concept of informed, concrete and conscious consent as the main basis of legitimization of personal data processing. Moreover, the technological possibilities of "Big Data" show that the depersonalization of personal data gradually ceases to be a guarantee of their anonymity.

In this regard, it is necessary either to radically reconsider the concept of personal data protection and to decide which data are goods and which are inalienable property benefits; or to create a higher quality technology capable of blocking deanonymization.

In any case, the business environment is forcing to expand the turnover of Big Data, the state, in response to the challenge of artificial intelligence, is seeking to make dot adjustments to the institution of personal information protection to prevent deanonymization, but under the conditions of enormous obvious and potential abilities of "Big Data", the preservation of personal data integrity seems more and more difficult and practically impossible.

References

- Brezgulevskaya, L. (2018). VKontakte & Double: case of banning the sale of data of social network database users. Available from: https://zakon.ru/blog/2018/7/12/v_kontakte_dabl_delo_o_zaprete_prodazhi_dannyh_polzovatelej_bazy_dannyh_socseti
- Federal Law of 27.07.2006. (2006). No152-FZ "On Personal Data". SPS "ConsultantPlus". Available from: <http://www.consultant.ru/>
- Rozhkova, M.A. (2019). What is big data, how they differ from ordinary data and what is the problem of legal regulation of big data. Available from: https://zakon.ru/blog/2019/4/22/chtoto_takoe_bolshie_dannye_big_data_chem_oni_otlichayutsya_ot_obychnyh_dannyh_i_v_chem_s_ostoit_proble
- Saveliev, A.I. (2015). Problems of personal data legislation application in the era of "Big Data". Law. Journal of the Higher School of Economics, 43–66.
- Sosnin K.A. (2019). Legal regulation of Big Date: foreign and domestic experience. Journal of the Court of Intellectual Rights, 30–43.
- Uroshleva, A. (2018). Commercialization of personal data and the concept of "big date" - topical issues of the IT-sphere. Garant.ru. Available from: <http://www.garant.ru/article/1229761/#ixzz64SV8SGZD>
- Volume of data/information created worldwide from 2005 to 2025 (in zettabytes). (2019). Statista. Available from: <https://www.statista.com/statistics/871513/worldwide-data-created/>
- Voynikanis, E.A., Semenova, E.V., Tyuliaev, G.S. (2018). Artificial intelligence and law: Challenges and opportunities of self-learning algorithms. Civil law. Civil process, 4, 137–148.

Citation:

Vasiliev, A.A., Pechatnova, Yu. V. (2020). Legal and environmental problems of personal data protection under commercialization of Big Data *Ukrainian Journal of Ecology*, 10(5), 133-135.



This work is licensed under a Creative Commons Attribution 4.0. License
