



**Анастасия Алексеевна
АНДРЕАСЯН,**

аспирант кафедры
информационного права
и цифровых технологий
Университета имени
О.Е. Кутафина (МГЮА)»,
заместитель руководителя
юридического отдела
ООО «Авиасейлс Медиа»
nastacia_1112@bk.ru
125993, Россия, г. Москва,
ул. Садовая-Кудринская, д. 9

Анализ подходов к определению больших данных в законодательстве Российской Федерации

Аннотация. В статье исследуются подходы к определению больших данных с точки зрения правового регулирования в Российской Федерации. Анализируются предлагаемые законодателем определения для включения в законодательство. Также исследуется разница в подходах к определению больших данных как массива информации и как технологии. Установлено, что в настоящее время термин «большие данные» имеет широкое употребление в различных областях и в зависимости от сферы и отрасли применения может меняться, что не способствует выработке единого подхода к регулированию. Отмечается, что имеющиеся ранее попытки дать определения большим данным в доктрине или на законодательном уровне приводили к описанию требований, предъявляемых к массивам данных или технологии в зависимости от отрасли рассмотрения.

Представляется, что такой подход является ограничивающим, не позволяет учесть всю сложность больших данных и может стать определенным барьером при применении в какой-то отдельной отрасли. Это будет требовать дополнительного разрешения на законодательном уровне, может привести к усложнению законодательства и, как следствие, сделать его неприменимым. Предлагается определять большие данные как сложный объект, состоящий из совокупности массива данных, характеризующихся разнообразием, и технологий, применимых для их обработки в режиме реального времени с использованием сети Интернет.

Ключевые слова: большие данные, данные, информация, информационно-телекоммуникационные сети, информационные технологии, правовая природа, правовое регулирование, сеть Интернет, техническое регулирование, цифровые технологии

DOI: 10.17803/2311-5998.2024.122.10.168-175

Anastasia A. ANDREASYAN,

Postgraduate student of the Department of Information
Law and Digital Technologies
of the Kutafin Moscow State Law University (MSAL),
Deputy Head of the Legal Department
Aviasales Media LLC
nastacia_1112@bk.ru
9, ul. Sadovaya-Kudrinskaya, Moscow, Russia, 125993

Analysis of approaches to the definition of big data in the legislation of the Russian Federation

Abstract. *The article examines approaches to the definition of big data from the perspective of legal regulation in the Russian Federation. It analyzes the definitions proposed by legislators for inclusion in legislation, as well as the difference in approaches to defining big data as an information array and as a technology. It is established that the term “big data” is currently widely used in various fields, and its meaning can vary depending on the specific industry or sector, which hinders the development of a unified regulatory approach. It has been noted that previous efforts to define big data at the doctrinal or legislative level have resulted in a description of the requirements for data or technologies, depending on their scope of application. This approach appears to be restrictive and fails to take into account the complex nature of big data, which may present a barrier when applied to certain industries, requiring additional legislative intervention. Such intervention could lead to an increase in the complexity of legislation, rendering it less applicable. Therefore, it is proposed to define big data as a complex entity consisting of a collection of diverse data sets, characterized by the technologies used to process them in real-time through the internet.*

Keywords: *big data, data, information, information and telecommunication networks, information technology, legal nature, legal regulation, Internet, technical regulation, digital technologies*

В абсолютном большинстве исследований в сфере правового регулирования в области цифровых технологий говорится о высокой ценности в настоящее время данных, которые называются новой нефтью¹. Одной из причин роста ценности данных является широкое распространение сети Интернет в масштабах всего мира и становление компаний IT-гигантов. Сегодня активно ставится вопрос о правовом регулировании оборота данных для большинства государств, так как оборот больших данных может быть полезен не только в рамках гражданского

¹ Андреасян А. А. Регулирование больших данных как элемента экосистемы // Индустрия 5.0, цифровая экономика и интеллектуальные экосистемы (ЭКОПРОМ-2021) : сборник трудов Всероссийской (Национальной) научно-практической конференции 18—20 ноября 2021 года. СПб. : Полтех-Пресс, 2021. С. 225—227.



оборота, но и для публичного управления или социальной сферы. Закономерным становится вопрос об определении больших данных в целом, поскольку анализ российского законодательства и доктрины показывает, что в настоящее время нет единого подхода к пониманию больших данных².

Было несколько попыток внедрить в российское законодательство определение больших данных, например, путем включения нового термина «большие пользовательские данные» в Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информационных технологиях и о защите информации» в 2018 г. Но законопроект в итоге не был принят. Тогда законодатель предлагал определять большие пользовательские данные как совокупность не содержащей персональных данных информации о физических лицах и (или) их поведении, не позволяющей без использования дополнительной информации и (или) дополнительной обработки определить конкретное физическое лицо, собираемой из различных источников (в том числе сети Интернет), количество которых превышает тысячу сетевых адресов³.

Законопроект критиковали за слишком поверхностное регулирование проблемы больших пользовательских данных из-за того, что он не давал ответа на вопрос, как соотносится предлагаемое регулирование с Федеральным законом от 27.07.2006 № 152-ФЗ «О персональных данных». Также важной признавалась следующая проблема. Предлагаемое определение больших пользовательских данных исходит из того, что пользовательские данные, не привязанные к конкретному сетевому адресу, не будут признаваться большими пользовательскими данными. Это противоречит подходу, который установился в ведущих мировых актах в области регулирования персональных данных.

Критике подверглось презюмирование прямой взаимосвязи массивов данных и Интернета. Несмотря на то, что развитие Интернета существенно повлияло на рост объемов данных, они могут быть получены не только с помощью сети Интернет⁴. Так, большие данные могут быть получены на научных установках класса «мегасайенс», например, на Большом адронном коллайдере⁵. Поэтому определение больших данных как массивов данных, получаемых с помощью сети Интернет, представляется ограничивающим, в то время как технология больших данных требует применения сети Интернет.

В 2020 г. Министерство цифрового развития, связи и массовых коммуникаций РФ выступило с инициативой внесения в Закон «Об информации, информационных технологиях и о защите информации» следующего определения: большие данные — совокупность неперсонифицированных данных, классифицируемую

² Полякова Т. А., Минбалеев А. В., Кроткова Н. В. Развитие доктрины российского информационного права в условиях перехода к экономике данных // Государство и право. 2023. № 9. С. 158—171.

³ Система обеспечения законодательной деятельности. URL: <https://sozd.duma.gov.ru/bill/571124-7> (дата обращения: 31.06.2024).

⁴ Савельев А. И. Направления регулирования больших данных и защита неприкосновенности частной жизни в новых экономических реалиях // Закон. 2018. № 5. С. 122—144.

⁵ Balyakin A., Mun D. Formation of an open science system in the European Union // Information and Innovation. 2017. No 1. P. 39—44.

по групповым признакам, в том числе информационные и статистические сообщения, сведения о местоположении движимых и недвижимых объектов, количественные и качественные характеристики видов деятельности, поведенческие аспекты движимых и недвижимых объектов, полученных от различных владельцев данных либо из различных структурированных или неструктурированных источников данных посредством сбора с использованием технологий, методов обработки данных, технических средств, обеспечивающих объединение указанной совокупности данных, ее повторное использование, систематическое обновление, форма представления которых не предполагает их отнесение к конкретному физическому лицу⁶. Данный законопроект также не был реализован.

Анализируя предложенный Министерством подход, можно отметить, что определение больших данных также дается через характеристику данных, они определяются как массив данных, который нельзя идентифицировать как персональные данные. Такой подход также представляется ограничивающим, так как он идет по пути определения объекта через то, чем он не является.

Одним из самых известных и популярных в литературе стал подход через определение больших данных с помощью нескольких V — как совокупность Volume, Velocity, Variety, где Volume означает объем, Velocity — скорость и Variety — разнообразие. Сейчас обсуждается и четвертое V — Value, означающее ценность. Данная дискуссия продолжается, и сейчас рассматривается расширение списка до включения Veracity, Variability, Visualization. То есть уже появляется требование о достоверности данных, их вариативности и возможности визуализировать их⁷. Предъявление дополнительных требований к большим данным, таких как достоверность, ценность и т.д., скорее, отвечает на запрос их практической применимости в той или иной сфере, что не умоляет значения данного определения. Вместе с тем нельзя сказать, что данные попытки дать определение больших данных отражают их суть в полном объеме.

Отметим, что все приведенные выше определения не учитывают, что большие данные являются также технологией, и рассматривают их исключительно как массив данных. Такой подход является одной из основ для определения больших данных как объекта интеллектуальных прав. Рассмотрение больших данных как базы данных фактически сразу исключает такой важный компонент, как технология. Этот подход также представляется ограничивающим.

Один из важнейших вопросов — как определять большие данные: как технологию или как массив данных⁸. Часто устанавливается технологический подход. Так, в Стратегии развития информационного общества в Российской Федерации

⁶ Законопроект о внесении изменений в Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информационных технологиях и о защите информации» // Федеральный портал проектов нормативных правовых актов. URL: <https://regulation.gov.ru/regulation/pna/publicview?npaid=99581> (дата обращения: 01.07.2024).

⁷ Интеллектуальное право в условиях развития технологии Big Data. База данных как объект интеллектуальных прав и иных прав : монография. М. : Проспект, 2022.

⁸ *Andreasyan A., Balyakin A., Nurbina M., Mukhamedzhanova A.* Security issues of scientific based Big Data circulation analysis // 8th International Conference on Data Science, Technology and Applications. 2019. P. 168—173.



на 2017—2030 годы, утвержденной Указом Президента РФ от 09.05.2017 № 203 обработка больших объемов данных указана как технология, являющаяся совокупностью подходов, инструментов и методов автоматической обработки структурированной и неструктурированной информации, поступающей из большого количества различных, в том числе разрозненных или слабосвязанных, источников информации, в объемах, которые невозможно обработать вручную за разумное время⁹. Более детального определения технологии больших данных на данный момент в российском законодательстве нет. Отчасти это можно связать с самой природой больших данных и изначально существующей дихотомией в подходе к их определению, а также с широким спектром их технологического применения.

Большие данные могут быть самостоятельной технологией, а могут использоваться для машинного обучения. Большие данные критически важны для обучения искусственного интеллекта в области медицины, так как позволяют сократить время на проведение расчетов при значимых для жизни операциях¹⁰ и могут помочь сохранить генетические банки данных, что является ресурсом для клинических исследований, т.е. научной инфраструктурой, используемой для открытий в науке¹¹.

Приведенные выше примеры показывают, что технология больших данных в зависимости от сферы ее применения находится на условном стыке отраслей, в том числе и отраслей права. Такое сложное положение приводит к тому, что правоприменитель вынужден подбирать и выбирать наиболее подходящее регулирование в каждой конкретной ситуации. Например, если вернуться к крупным IT-компаниям, то обладание ими большими данными фактически позволяет им занимать доминирующее положение на рынке и ограничивает появление новых игроков, так как они не способны сформировать аналогичный ресурс данных, и в таком случае применяется антимонопольное регулирование. Ряд стран в настоящее время разрабатывают и принимают акты, направленные на сдерживание и установление рамок для таких участников рынка¹².

Актуальность вопроса о правовом регулировании больших данных подтверждается тем, что Национальная программа «Цифровая экономика» указывает на необходимость создания специальных правовых режимов, позволяющих

⁹ Указ Президента РФ от 09.05.2017 № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017—2030 годы // URL <http://www.kremlin.ru/acts/bank/41919> (дата обращения: 01.07.2024).

¹⁰ *Isaev A., Dobroserdova T., Danilov A., Simakov S.* Physically Informed Deep Learning Technique for Estimating Blood Flow Parameters in Four-Vessel Junction after the Fontan Procedure // 10th Anniversary of Computation—Computational Biology. URL: <https://www.mdpi.com/2079-3197/12/3/41> (дата обращения: 01.07.2024).

¹¹ *Кнопперс Б. М.* Генетика, геномика и права человека // Прецеденты Европейского суда по правам человека. 2018. № 3. С. 35—41.

¹² К таким актам можно отнести акт The Digital Service Act EU (2022), целью которого является дать наибольший доступ компаниям к возможности размещения своих приложений во внутренних магазинах и снизить ограничения крупных IT-компаний (URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en (дата обращения: 04.07.2024)).

обеспечить беспрепятственное развитие сквозных технологий, к которым также отнесены большие данные. В 2020 г. был принят Федеральный закон № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации», который предусматривает возможность для государства делать изъятия из регулирования для предоставления компаниям возможности использовать и развивать на практике свои инновационные разработки на основе сквозных технологий.

Данные решения со стороны государства лишь подтверждают достаточно популярный среди исследователей в области регулирования цифровых технологий тезис о том, что технологии развиваются стремительно и право фактически вынуждено успевать за этим развитием¹³.

Представляется, что одна из сложностей в выработке определения, отражающего сущность сложного объекта информационно-правовых отношений, состоит в отсутствии единого подхода к его пониманию¹⁴. В том числе использование термина «большие данные» в различных значениях (от технического значения, например, в международных стандартах ИСО до доктринальной литературы) подтверждает то, что большие данные сейчас, скорее, технический термин, чем правовой.

Фактически большие данные являются сложным объектом и рассмотренные выше способы определять большие данные показывают, что существуют значимые требования к массиву данных, что подтверждается ростом условных V в их определении (ценность, возможность визуализации, объем, достоверность и т.д.). В то же время расширение требований к большим данным как массиву данных может свидетельствовать об их динамичном развитии и гибкости. Если применять расширительное толкование, то саму информацию можно понимать как любые информационные объекты, в том числе как определенную совокупность данных, связанных уникальным набором технологий, с помощью которых происходит их обработка¹⁵.

При этом большие данные также являются технологией, что отражается в ряде отраслевых актов. Например, в 2021 г. был утвержден Национальный стандарт РФ ГОСТ Р ИСО/МЭК 20546-2021 «Информационные технологии. Большие данные. Обзор и словарь»¹⁶, который определяет большие данные как большие массивы данных, отличающиеся главным образом такими характеристиками, как объем, разнообразие, скорость обработки и/или вариативность, что требует использования технологии масштабирования для эффективного хранения, обработки, управления и анализа.

¹³ Минбалеев А. В. Проблемы регулирования искусственного интеллекта // Вестник Южно-Уральского государственного университета. Серия : Право. 2018. № 4. С. 82—87.

¹⁴ Лескина Э. И. Технологии больших данных (big data) для охраны окружающей среды // Экологическое право. 2023. № 1. С. 21—24.

¹⁵ Минбалеев А. В. Система информации: теоретико-правовой анализ : дис. ... канд. юрид. наук. Челябинск, 2006.

¹⁶ Большие данные (Big Data) // URL: <https://protect.gost.ru/document1.aspx?control=31&baseC=6&page=4&month=4&year=-1&search=&id=240981> (дата обращения: 04.07.2024).



Исключить совсем такое многообразие использования термина не представляется возможным, но в то же время данное обстоятельство все больше подчеркивает важность единого понимания больших данных с точки зрения правового регулирования.

Учитывая изложенное, предлагается определять большие данные как сложный объект, состоящий из совокупности массива данных, характеризующихся разнообразием, и технологий, применимых для их обработки в режиме реального времени с использованием сети Интернет.

При выборе подхода к определению больших данных следует учитывать разнообразие сфер их применения — от медицины и большой науки до коммерческого использования и публичного управления. Важно, чтобы выбранное регулирование не стало фактором, препятствующим внедрению и развитию технологии. Представляется, что предложенное определение позволит учесть сложность исследуемого объекта.

БИБЛИОГРАФИЯ

1. *Андреасян А. А.* Регулирование больших данных как элемента экосистемы // Индустрия 5.0, цифровая экономика и интеллектуальные экосистемы (ЭКО-ПРОМ-2021) : сборник трудов Всероссийской (Национальной) научно-практической конференции 18—20 ноября 2021 г. — СПб. : Полтех-Пресс, 2021. — С. 225—227.
2. Интеллектуальное право в условиях развития технологии Big Data. База данных как объект интеллектуальных прав и иных прав / Е. А. Войниканис, М. А. Кольздроф, В. А. Корнеев, Е. В. Ульянова, Н. А. Шибанова. — М. : Проспект, 2022. — 176 с.
3. *Кнопперс Б. М.* Генетика, геномика и права человека // Прецеденты Европейского суда по правам человека. — 2018. — № 3. — С. 35—41.
4. *Лескина Э. И.* Технологии больших данных (big data) для охраны окружающей среды // Экологическое право. — 2023. — № 1. — С. 21—24.
5. *Минбалеев А. В.* Проблемы регулирования искусственного интеллекта // Вестник Южно-Уральского государственного университета. — Серия : Право. — 2018. — № 4. — С. 82—87.
6. *Минбалеев А. В.* Система информации: теоретико-правовой анализ : дис. ... канд. юрид. наук. — Челябинск, 2006. — 272 с.
7. *Полякова Т. А., Минбалеев А. В., Кроткова Н. В.* Развитие доктрины российского информационного права в условиях перехода к экономике данных // Государство и право. — 2023. — № 9. — С. 158—171.
8. *Савельев А. И.* Направления регулирования больших данных и защита неприкосновенности частной жизни в новых экономических реалиях // Закон. — 2018. — № 5. — С. 122—144.
9. *Andreasyan A., Balyakin A., Nurbina M., Mukhamedzhanova A.* Security issues of scientific based Big Data circulation analysis // 8th International Conference on Data Science, Technology and Applications. — 2019. — P. 168—173.

10. *Balyakin A., Mun D.* Formation of an open science system in the European Union // Information and Innovation. — 2017. — No. 1. — P. 39—44.
11. *Isaev A., Dobroserdova T., Danilov A., Simakov S.* Physically Informed Deep Learning Technique for Estimating Blood Flow Parameters in Four-Vessel Junction after the Fontan Procedure // 10th Anniversary of Computation—Computational Biology. — URL: <https://www.mdpi.com/2079-3197/12/3/41> (дата обращения: 04.07.2024).

