

## БОЛЬШИЕ ДАННЫЕ: ПРАВОВОЙ АСПЕКТ

М. С. Абламейко

*доцент кафедры конституционного права  
юридического факультета Белорусского государственного университета,  
кандидат юридических наук, доцент  
m.ablameyko@mail.ru*

В статье рассматриваются подходы к правовому регулированию института больших данных с учетом международного опыта.

**Ключевые слова:** большие данные; искусственный интеллект; цифровая трансформация, цифровизация; цифровое законодательство.

Развитие информационного общества, предоставление всеобщего доступа к сети Интернет, цифровая трансформация всех сфер жизнедеятельности общества в целом, привели к росту предоставляемой и потребляемой информации. В настоящее время информация стала настолько ценным ресурсом, что сложно преувеличить ее значение. Знаменитая фраза Натана Майера Ротшильда «Кто владеет информацией, тот владеет миром» как никогда отражает реальность. Количество собираемых, обрабатываемых, генерируемых и т.д. данных растет в геометрической прогрессии. Это связано, в том числе, с переходом от аналогового формата на электронный (цифровой), что позволяет не только хранить огромное количество информации, но и использовать ее в различных целях. По прогнозам ООН объем данных на глобальном уровне достигнет 175 зеттабайт в 2025 г. [1].

Таким образом, в настоящее время формируются подходы к регулированию института больших данных. Сложность заключается в том, что большие данные, в отличие от открытых, могут включать в себя информацию ограниченного доступа. В связи с этим необходимо сформировать комплексный подход в регулировании данного вопроса, чтобы при сохранении приоритета развития технологий создать необходимые условия для защиты интересов личности, общества и государства и обеспечить национальную и информационную безопасность.

В юридической науке обсуждаются вопросы анализа правовых последствий использования больших данных, включая соблюдение норм, касающихся приватности и защиты персональных данных, кибербезопасности и антимонопольного права (использование данных крупными корпорациями). Акцент делается на правовых вызовах и принятии мер по регулированию.

К вопросу как большие данные меняют общество, экономику и право обращались Виктор Майер-Шенбергер и Кеннет Кукеэр в книге «Big Data: A Revolution That Will Transform How We Live, Work, and Think». Авторы подчеркивают необходимость правовых рамок для регулирования сбора и использования данных, особенно в контексте приватности и предсказательной аналитики [2]. Лучиано Флориди рассматривает большие данные через призму этики, с точки зрения морально обоснованных решений и прозрачности алгоритмов [3]. Риски, связанные с большими данными, и возникающие

угрозы конфиденциальности, такие как слежка, раскрытие информации и дискриминация анализирует американский юрист Райан Кало [4]. Деятельность платформ, таких как Google, Amazon и др., использование ими данных с целью доминирования на рынке рассматривает Аннабель Гавер.

На данном этапе правовое закрепление понятия «больших данных» как на международном, так и на национальном уровнях отсутствует. Его аспекты регулируются через нормы, связанные с обработкой информации, защитой данных и цифровизацией. В законодательных актах Российской Федерации, Кыргызской Республики, Республики Армения в сфере информации и информационных технологий термин информация определен через категорию данные. В Республике Беларусь и Республике Казахстан понятие информации определяется через сведения. Следует отметить, что в Республике Казахстан, в отличии от других стран, в Административном процедурно-процессуальном кодексе от 29 июня 2020 г. № 350-VI разъясняется понятие данных как информации в формализованном виде, пригодной для обработки.

В настоящее время термин «большие данные» определен в актах технического регулирования. В соответствии с Национальным стандартом Российской Федерации ГОСТ Р ИСО/МЭК 20546-2021 «Национальный стандарт Российской Федерации. Информационные технологии. Большие данные. Обзор и словарь», идентичным Международному стандарту ИСО/МЭК (ISO/IEC) 20546:2019 «Information technology – Big data – Overview and vocabulary», IDT) большие данные (big data) – большие массивы данных, отличающиеся главным образом такими характеристиками, как объем, разнообразие, скорость обработки и/или вариативность, которые требуют использования технологии масштабирования для эффективного хранения, обработки, управления и анализа. Также отмечено, что термин «большие данные» широко применяется в различных значениях, например в качестве наименования технологии масштабирования, используемой для обработки больших массивов данных.

Большие данные можно рассматривать не только как большие массивы данных, но и как технологию. В связи с тем, что такого рода технологии пронизывают все сферы жизнедеятельности и их принято считать сквозными, то и правовое регулирование должно быть комплексным.

Классификация больших данных формировалась постепенно, в соответствии с потребностями анализа и обработки растущих объемов информации в эпоху цифровизации.

Модель 3V: volume, velocity, variety, предложенная Дугласом Лэйни аналитиком компании Gartner, легла в основу современной классификации больших данных. В рамках данной концепции рассматриваются такие характеристики больших данных как: объем – количество данных, измеряющееся в гигабайтах, зеттабайтах или йоттабайтах; скорость – определяет насколько быстро массивы информации генерируются и обрабатываются; разнообразие – различные типы данных, которые собираются из разных источников, что определяет форматы таких данных (структурированные, неструктурированные и др.) [5].

Позже, усилиями экспертов, компаний и исследовательских сообществ в связи с развитием технологий и возрастающими потребностями рынка, к данной модели добавились другие критерии: veracity (достоверность) – качество данных, позволяющее повысить

точность результатов и value (ценность) – возможность выявления скрытых закономерностей, прогнозирования, оптимизации процессов, персонализации предложений, т.е выгода, которую организация получает от данных, что сформировало концепт 5V.

В настоящее время классификация данных продолжает расширяться. Выделяются дополнительные характеристики: volatility (изменчивость) – данные подвержены постоянному изменению; visualization (визуализация) – делает большие данные доступными для человеческого восприятия; validity (валидность) – процесс проверки данных различных типов по критериям корректности и полезности для конкретного применения; vulnerability (уязвимость) и viscosity (сложность обработки данных) и др. В связи с постоянным ростом количества данных, а также скорости их обработки увеличивается их разнообразие. Выделяются структурированные, полуструктурные и неструктурные. По отраслевой принадлежности большие данные можно разделить на: промышленные, финансовые, аграрные, экологические, демографические и др. В зависимости от владельца: публичные (государственные); корпоративные (принадлежащие определенной организации); частные.

Особенно актуальным вопрос регулирования больших данных становится в связи с развитием технологий ИИ. Это связано с тем, что на этапе тренировки нейросетевых структур, требуются большие объемы размеченных данных для различных прикладных задач, причем важно обучить систему ИИ именно на реальных больших данных. Регулирование технологий больших данных, как и технологий ИИ, возможно в рамках определенных отраслей права в зависимости от сферы применения.

Международное сообщество стремится выработать общие подходы к регулированию больших данных таким образом, чтобы удалось соблюсти баланс стимулирования развития технологий, с одной стороны, и безопасность их функционирования, с другой. Решение данной задачи возможно с технической и юридической точки зрения. Внедрение технологий ИИ в общественные процессы не должно приводить к ущемлению прав и свобод человека независимо от экономического эффекта и коммерческого интереса.

Первыми шагами в области регулирования данного вопроса, с целью поиска баланса между стимулированием развития технологий и защитой прав и свобод личности и общества стало определение этических принципов и норм.

Значимым документом являются Рекомендации по этике ИИ (Recommendation on the Ethics of Artificial Intelligence) [6], которые позиционируется ЮНЕСКО как инструмент «мягкого права» и не носят обязательного характера. Одним из стратегических направлений реализации Рекомендаций является – политика в области данных. В первую очередь государствам-членам рекомендуется разработать стратегии управления данными, обеспечивающие непрерывную оценку качества данных, используемых для обучения систем ИИ, включая адекватность процессов сбора и выборки данных, надежность мер по обеспечению безопасности и защиты данных, включая соответствующие гарантии соблюдения права на неприкосновенность личной информации на всем протяжении жизненного цикла систем ИИ [6]. Например, Правительство Великобритании приняло Стратегию по данным Великобритании (National Data Strategy) для обмена данными между всеми игроками рынка (B2B, B2G, G2B, G2G) как на национальном, так и на международном уровнях для стремительного развития новых технологий; в Германии утверждена Стратегия по данным Федерального Правительства Германии (Data Strategy of the Federal German Government), целями которой являются: создание

эффективной и устойчивой инфраструктуры данных в стране, повышение инновационного и ответственного использования данных, развитие компетентности в области данных и формирование культуры данных; в рамках правительственные программы по цифровому развитию Индии Digital India разработана инициатива «Открытые данные» (Open Government Data (OGD) Platform) [5]. Также, государствам-членам следует разработать правила и стандарты в отношении данных или эквивалентные рамочные принципы, либо провести пересмотр существующих положений для обеспечения защиты персональных и конфиденциальных данных, которые, в случае их разглашения, могут причинить людям значительный ущерб, физический вред или страдания (данные о совершении правонарушений, биометрические и генетические данные, данные о здоровье и др.) [6].

Таким образом, для формирования политики в области больших данных необходимо принять стратегию, в которой отразить юридические, технические и организационные меры, регулирующие данный институт. На законодательном уровне следует решить ряд вопросов, включая: определение терминологии, правила предоставления доступа к большим государственным структурам; принципы классификации данных и разграничение больших данных от персональных данных путем определения методов обезличивания персональных данных; обеспечение защиты прав и свобод человека в цифровом пространстве, включая неприкосновенность частной жизни. Важным этапом является формирование стандартов в данной области для защиты информации конфиденциального характера и персональных данных. На отраслевом уровне следует предусмотреть возможность использования данных конфиденциального характера только после прохождения обезличивания (анонимизации).

### **Библиографический список**

1. Инновации в сфере данных для целей развития [Электронный ресурс]. – URL: <https://www.un.org/ru/global-issues/big-data-for-sustainable-development>. (дата обращения: 23.04.2025).
2. Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / В. Майер-Шенбергер, К. Кукеэр ; пер. с англ. Инны Гайдюк. – М. : Манн, Иванов и Фербер, 2014. – 240 с.
3. Floridi, L., Taddeo, M. What is Data Ethics? / Philosophical Transactions of The Royal Society a Mathematical Physical and Engineering Sciences. Vol. 374. No. 2083. 2016. Art. 20160360. – URL: <https://doi.org/10.1098/rsta.2016.0360>. – (date of access: 23.04.2025).
4. Calo, R. Consumer Subject Review Boards: A Thought Experiment: forum the center for internet and society, Stanford Law School, 3 Sept. 2013 / Stanford Law Review Online 97-102. – URL: – Режим доступа: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2340745](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2340745)– (date of access: 23.04.2025).
5. Laney, D. 3D Data Management: Controlling Data Volume, Velocity and Variety / Application Delivery Strategies, Meta Group, 2001. – 3 р.
6. UNESDOC [Цифровая библиотека] – URL: [https://unesdoc.unesco.org/ark:/48223/pf000380455\\_rus](https://unesdoc.unesco.org/ark:/48223/pf000380455_rus) (date of access: 13.03.2025).
7. Глобальный атлас регулирования искусственного интеллекта. Издание 2023 / Коллектив авторов под редакцией А. В. Незнамова. – М., 2023. – 452 с.