

Проблемы соотношения персональных данных и больших данных

Научный руководитель – Зайнутдинова Елизавета Владимировна

Сизикова Софья Алексеевна

Студент (бакалавр)

Новосибирский национальный исследовательский государственный университет,

Новосибирск, Россия

E-mail: sizikova.sofa03@gmail.com

В условиях цифровизации активное развитие получили технологии перевода информации в электронный вид и работы с большими массивами информации. Одним из ключевых терминов в данной области стали «большие данные» (Big Data), которые активно используются для обучения искусственного интеллекта.

Прежде всего стоит определить сущность больших данных. В зарубежной литературе распространение получил подход 3V: объём (volume), скорость (velocity) и разнообразие (variety) [6]. Таким образом, большие данные являются совокупностью огромных массивов различной информации, которые подвергаются дальнейшей обработке, которая происходит с помощью специальных технологий, которые неотделимы от понятия больших данных. Некоторые исследователи отмечают существование двух видов больших данных: промышленные – показатели оборудования (температуры, давления и пр.) и пользовательские – данные от человека (например, его цифровой след) [4]. Таким образом, логически формируется вывод о том, что персональные данные являются разновидностью больших данных. Однако представляется, что большие данные правомерно могут включать в себя персональные данные только в обобщенном и обработанном виде, то есть с потерей возможности соотнесения таких сведений с конкретным лицом.

Стоит обратить внимание, что персональные данные, полученные из открытых источников, являются неотъемлемой частью больших данных и позволяют бизнесу значительно повышать качество принимаемых управленческих решений на основе детального анализа поведения потребителей. Большое значение в определении правомерности обработки данных пользователей из социальных сетей имеет дело ООО «В Контакте» (далее – Вконтакте) против ООО «Дабл» (далее – Дабл). В этом споре разрешался вопрос законности использования базы данных, однако его судебное разрешение косвенно определило бы судьбу парсинга, то есть сбора и систематизации информации из открытых интернет-источников (фактически одна из составляющих работы с большими данными). Судом первой инстанции в иске Вконтакте отказано, судом апелляционной инстанции решение первой инстанции отменено. Апелляция указала, что по лицензионному соглашению Вконтакте, воспроизведение, копирование, сбор, систематизация, хранение, передача информации из Вконтакте в коммерческих целях или в целях извлечения базы данных или ее использования полностью или в любой части не допускается [2]. Следовательно сбор данных пользователей и дальнейшее их распространение без согласия субъектов персональных данных неправомерны. Судом кассационной инстанции производство по делу прекращено в связи с заключением мирового соглашения, что оставило без ответа вопрос о правомерности парсинга персональных данных из открытых источников. Представляется, что сбор и систематизация общедоступной информации из социальных сетей все-таки должны допускаться для развития бизнеса, улучшения клиентского сервиса, однако недопустимо распространение таких сведений без согласия пользователей, а также использование персональных данных как отдельного объекта в коммерческих целях.

Стоит обратить внимание на адресность согласия на обработку персональных данных [1]. Отмечается также, что если граждане дали согласие на обработку их персональных данных администратору доменного имени «ВКонтакте», который разместил данные сведения на своем ресурсе, то это не обуславливает право других лиц на их свободное использование, распространение.

В этом же ключе может быть упомянут принцип соответствия обработки персональных данных ее цели, указанной в согласии, в ином подходе принцип недопустимости объединения баз данных, собранных с несовместимыми целями. С целью соблюдения данных принципов существует необходимость получения согласия субъекта персональных данных, которые включаются в большие данные, в случае использования сведений в коммерческих целях [4].

Высказывается позиция об абсолютной недопустимости оборота персональных данных, так как даже по обезличенным персональным данным, с которых в принципе снимается режим конфиденциальности, иногда можно восстановить личность, например, по списку покупок можно узнать регион проживания, изменение семейного статуса или здоровья человека, иногда даже опознать личность [5]. Представляется, что такой подход может значительно замедлить развитие технологий искусственного интеллекта, так как тренировка ИИ производится с помощью больших данных. Разработка специального федерального закона о больших данных избыточна, однако разграничение больших и персональных данных необходимо.

Таким образом, дополнительному внесению ясности в проблему разграничения больших и персональных данных может способствовать закрепление в постановлении Пленума Верховного Суда Российской Федерации, к примеру: «В работе с технологиями, используемыми для обучения искусственного интеллекта, возможно использование персональных данных исключительно в анонимизированном виде, в противном случае на такие сведения распространяются правила, установленные Законом о персональных данных. Не допускается использование персональных данных без согласия субъекта таких данных».

Источники и литература

- 1) Определение Второго кассационного суда общей юрисдикции от 13 июля 2021 г. по делу № 88-15635/2021, 2-1712/2020: Справочная правовая система «Консультант Плюс».
- 2) Постановление Девятого арбитражного апелляционного суда № 09АП-31545/2021-ГК от 08 июля 2021 г. по делу № А40-18827/2017: <https://kad.arbitr.ru/>.
- 3) Савельев А.И. Направления регулирования Больших данных и защита неприкосновенности частной жизни в новых экономических реалиях // Закон. 2018. № 5. С. 122-144.
- 4) Савельев А.И. Проблемы применения законодательства о персональных данных в эпоху «больших данных» // Право. Журнал Высшей школы экономики. 2015. № 1. С. 43-66.
- 5) Солдаткина О.Л. Проблемы сбора больших данных с позиции законодательства о персональных данных // Устойчивое развитие России: правовое измерение: сб. докладов X Московского юридического форума. В 3-х частях. М., 2023. С. 178-180.
- 6) Jain, P., Gyanchandani, M. & Khare, N. Big data privacy: a technological perspective and review // Journal of Big Data № 3. 2016, Article № 25.