УДК 004.056.5

**Проблемы анализа данных и регламент GDPR**

**Аннотация**

В современных условиях очевидное влияние на развитие отрасли связи, в том числе на развитие интернет- провайдеров в России оказывает значительное влияние активное развитие законотворчество в области сбора, хранения и переработки данных пользователей услуг связи и прежде всего – телекоммуникационных услуг, в том числе передачи данных в глобальных сетях.

Данная статья направлена на анализ возможных последствий вступления в силу в мае 2018 года Регламента General Data Protection Regulation (GDPR) Европейского Союза для отечественных операторов персональных данных (интернет-компаний и других телекоммуникационных компаний) предоставляющих услуги через глобальную сеть для клиентов в странах ЕС, в контексте действующего и разрабатываемых систем регулирования в Российской Федерации.

Ключевые слова: большие данные, защита ПДн, GDPR, конфиденциальность, анализ данных

**Введение**

Термин большие данные описывает большие или сложные объёмы данных, как структурированные, так и неструктурированные, в которых можно анализировать, такие параметры, как объем, скорость и разнообразие[1]. Сегодня большие данные стали капиталом, причем предприятия существенно улучшают свою деятельность и отношения с клиентами, а научные круги развивают и расширяют исследования (например, в области климата [3] или биологических исследований [2]). Кроме того, для огромного объема, скорости генерации и разнообразия данных требуются специальные архитектуры для хранения и обработки (например, MapReduce [4] или Apache Hive1). С юридической точки зрения, в этом документе все фокусируется на общем регламенте ЕС по защите данных (GDPR) [5], который вступил в силу в мае 2018 года. Он актуален для всех организаций внутри Европейского Союза (ЕС), Европейского Экономическое пространство (EEA), а также организаций из других стран, если они обрабатывают данные европейских граждан. Таким образом, GDPR оказывает влияние на большинство крупных компаний во всем мире.

**Общий обзор**

GDPR регулирует сбор, хранение и обработку персональных данных. К персональным данным относятся любые данные, которые могут быть связаны с конкретным физическим лицом. Сюда входят не только прямые персональные идентификаторы (например, полное имя, национальный идентификационный номер), но и косвенные идентификаторы. например, номера телефонов, IP-адреса или фотографии с идентифицируемыми людьми. Данные, которые не включают такие идентификаторы, обычно считаются анонимными и выходят за рамки GDPR. Результаты анализа больших данных очень часто представляют собой безссылочные статистические данные, поэтому самый простой способ соответствовать всем требованиям GDPR - это обрабатывать только анонимные данные. Однако определение анонимности не является тривиальным. Даже если из набора данных удаляют идентифицируемые параметры, существует ненулевая возможность повторной идентификации отдельных лиц, путём объединения набора данных с другой информацией. Этот подход для деанонимизации называется атакой с помощью фоновых знаний [6], [7].

**GDPR как европейский регулятор защиты ПДн**

Цель внедрения новой редакции GDPR в систему защиты персональных данных состоит в том, чтобы защитить всех граждан ЕС от нарушения конфиденциальности и нарушениях в процессе доступа к данным в современном мире, который управляется данными и их производными. Хотя основные принципы конфиденциальности данных по-прежнему остаются такими же, как и в предыдущей директиве, в новой редакции в политику регулирования было предложено внести множество изменений, где ключевые из них:

Экстерриториальность. Возможно, самое большое изменение в регулирующем поле конфиденциальности данных связано с расширенной юрисдикцией GDPR, поскольку она применяется ко всем компаниям, обрабатывающим персональные данные субъектов данных, проживающих в Евросоюзе, независимо от места прописки компании, когда ее деятельность связана с предложением товаров или услуг гражданам ЕС (независимо от того, требуется ли оплата) и мониторингом поведения, которое имеет место в рамках ЕС. Не входящие в ЕС предприятия, обрабатывающие данные граждан ЕС, должны назначить своего представителя в ЕС.

Важно отметить, что эти правила применяются как к контроллерам, так и к процессорам – то есть так называемые "облака" не освобождаются от применения GDPR.

Согласие. Теперь условия для получения согласия были усилены и компании больше не могут использовать длинные неразборчивые условия мелким шрифтом. Согласие должно быть чётким и отличимым от других вопросов и предоставляться в понятной и легкодоступной форме, используя ясный и простой язык. Отозвать согласие должно быть так же легко, как и дать его.​

## Права субъекта данных. Уведомление о нарушении в соответствии с GDPR, уведомления о нарушении теперь являются обязательными во всех государствах-членах, где нарушение данных может “привести к риску для прав и свобод отдельных лиц”. Это должно быть сделано в течение 72 часов после того, как впервые стало известно о нарушении. Обработчики данных также обязаны уведомлять своих клиентов“без неоправданной задержки” после того, как впервые стало известно о нарушении данных.

Право на доступ. Частью расширенных прав субъектов данных, изложенных в GDPR, является право субъектов данных получать подтверждение от контроллера данных о том, обрабатываются ли персональные данные, касающиеся их, где и с какой целью.

Кроме того, контролёр должен бесплатно предоставить копию персональных данных в электронном формате. Это изменение представляет собой резкий сдвиг в сторону транспарентности данных и расширения возможностей субъектов данных.

Контроллер данных (организация, ответственная за сбор данных) не может ни определить слишком общую цель обработки данных, ни изменить её позже произвольно ( Рис. 1).

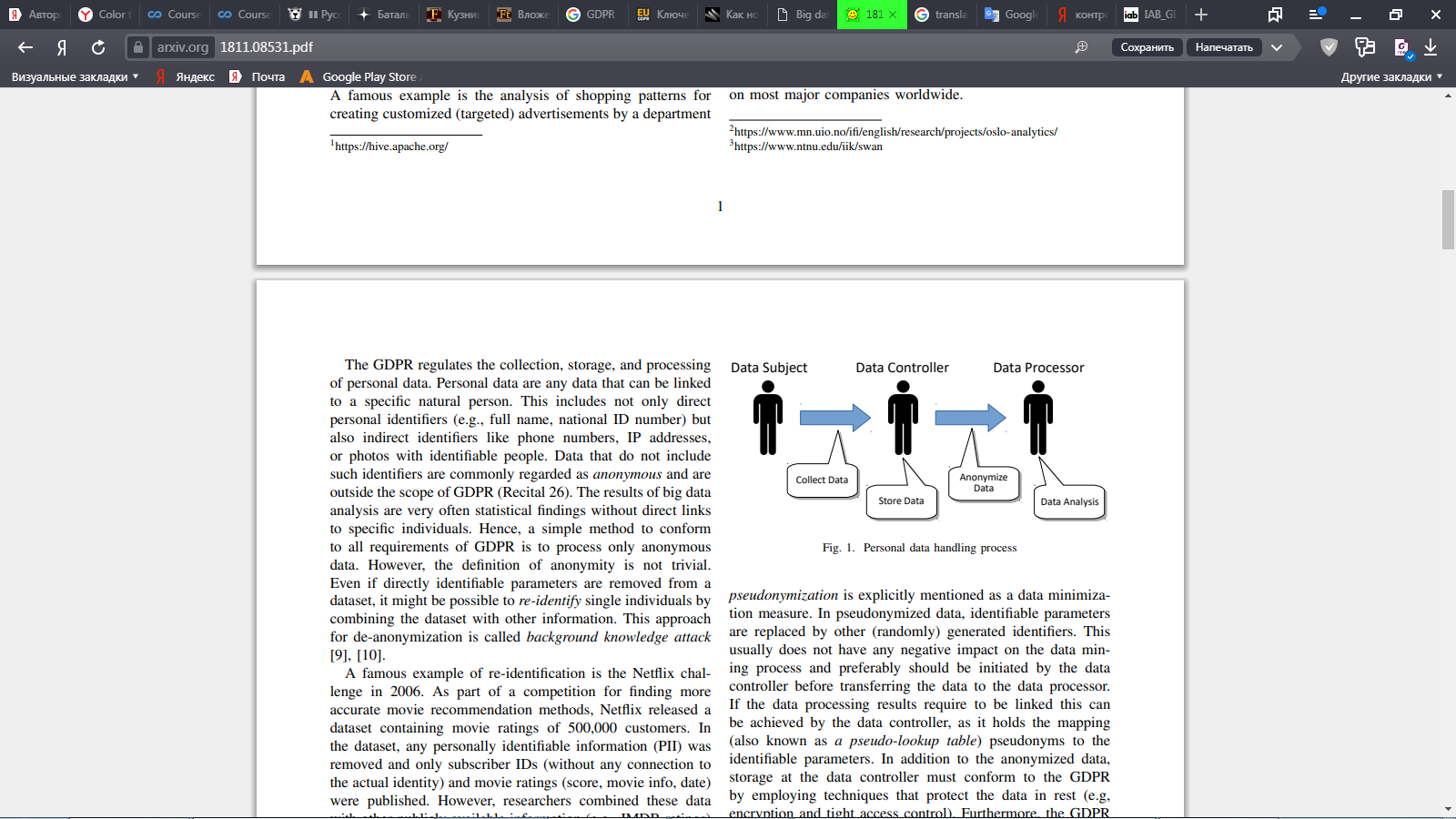


Рис. 1. Процесс обработки персональных данных

Право быть забытым. Также известное как стирание данных, право быть забытым дает субъекту данных право заставить контроллера данных стереть его / ее личные данные, прекратить дальнейшее распространение данных и предожить третьим лицам прекратить обработку данных. Условия для стирания, изложенные в статье 17, включают данные, которые более не имеют отношения к первоначальным целям обработки, или отзыв согласия субъекта данных.

Переносимость данных. GDPR вводит переносимость данных- право субъекта данных получать персональные данные, относящиеся к ним, – которые они ранее предоставляли в "широко используемом и машиночитаемом формате" и имеют право передавать эти данные другому контроллеру.

Конфиденциальность. "контроллер должен ... эффективно осуществлять соответствующие технические и организационные меры ... в целях выполнения требований настоящего Положения и защиты прав субъектов данных". Статья 23 предусматривает, что контроллеры должны хранить и обрабатывать только те данные, которые абсолютно необходимы для выполнения их обязанностей (минимизация данных), а также ограничивать доступ к персональным данным тем, кто должен осуществлять обработку.

## Сотрудники по защите данных. В соответствии с GDPR нет необходимости представлять уведомления /регистрации каждому местному оператору по защите данных о деятельности по обработке данных, а также не требуется уведомлять/ получать одобрение на передачу данных на основе положений типового контракта. Вместо этого, существуют внутренние требования к ведению учёта и назначение специалиста ответственного по защите данных (DPO), что является обязательным только для тех контроллеров и операторов, основная деятельность которых состоит из операций по обработке, требующих регулярного и систематического мониторинга субъектов данных в крупном масштабе или специальных категорий данных или данных, касающихся уголовных обвинений и преступлений. Важно отметить, что данные физических лиц могут быть связаны с - идентификаторами, предоставляемыми их устройствами, приложениями, инструментами и протоколами, такими как адреса интернет- протокола, идентификаторы файлов cookie или другие идентификаторы, такие как теги радиочастотной идентификации. Это может оставить следы, которые, в частности, в сочетании с уникальными идентификаторами, полученной серверами, могут использоваться для создания профилей физических лиц и их идентификации.

## Учитывая четкую позицию экстерриториальности регламента ЕС, российским операторам связи и организаторам распространения информации для минимизации возможных рисков потребуется:

## 1. Осуществить аудит заключенных договоров с компаниями из государств-членов Евросоюза для приведения в соответствие условий договоров с требованиями Регламента GDPR, а также с учетом императивных требований норм Пакета Яровой;

## 2. Внести соответствующие изменения в действующие договоры с компаниями-контрагентами из стран Европейского Союза и распределить возможные финансовые затраты;

## 3. Назначить в рамках компании компетентное лицо (группу лиц), ответственных за мониторинг разъясняющих норм применения Регламента GDPR, включая подзаконные и имплементирующие акты, принятые в соответствии с настоящим Регламентом и с правом государств-членов Евросоюза для уточнения положений Регламента GDPR, которые в перспективе могут быть приняты государствами членами Евросоюза и Комиссией Европейского Союза;

## 4. Назначить своего представителя в соответствующем государстве члене Евросоюза для взаимодействия с национальными надзорными органами.

## **Заключение**

## В контексте больших данных нельзя исключать, что анализ данных может касаться "конфиденциальных данных", обработка которых в большинстве случаев ограничена и запрещена, или что он будет иметь “трансформационное воздействие” на данные. Например, обработка несекретных персональных данных может привести – через интеллектуальный анализ данных, например, – к созданию данных, которые раскрывают конфиденциальную информацию о человеке.

Широкая сфера применения GDPR и возможная обработка конфиденциальных данных могут потребовать в рамках ЕС ограничения определенных видов деятельности по обработке или технических разработок для решения жестких правил, включенных в GDPR.

**Список использованной литературы:**

1. Грушка Н. и др. Вопросы конфиденциальности и защиты данных в больших данных: анализ кейсов в рамках GDPR //международная конференция IEEE 2018 по большим данным (Big Data). - IEEE, 2018. – С. 5027-5033.
2. Маркс, Вивьен. "Биология: большие проблемы больших данных." (2013): 255. [Online]. Available: https://www.nature.com/articles/498255a

3. S. Bony, J. T. Overpeck, G. A. Meehl and D. R. Easterling, “Climate Data Challenges in the 21st Century,” Science, vol. 331, no. 6018, pp. 700–702, Feb. 2011.

4. J. Dean and S. Ghemawat, “MapReduce: simplified data processing on large clusters,” Communications of the ACM, vol. 51, no. 1, pp. 107–113, 2008. [Online]. Available: https://www.nature.com/articles/ejhg2015239

5.European Parliament and Council, “Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance),” May 2016. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/oj/eng

6. D. Kifer and A. Machanavajjhala, “No free lunch in data privacy,” in Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, ser. SIGMOD. New York, NY, USA: ACM, 2011, pp. 193–204. [Online]. Available: [http://doi.acm.org/10. 1145/1989323.1989345](http://doi.acm.org/10.%201145/1989323.1989345)

7. A. Machanavajjhala, M. Venkitasubramaniam, D. Kifer, and J. Gehrke, “L-Diversity: Privacy Beyond k-Anonymity,” in 22nd International Conference on Data Engineering (ICDE’06)(ICDE), 2006, p. 24. [Online]. Available: doi.ieeecomputersociety.org/10.1109/ICDE.2006.1

© В.М.Горяев, О.Ю.Пескова, 2019