Доверие и безопасность в информационном обществе

ДОВЕРИЕ И БЕЗОПАСНОСТЬ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ В РОССИИ

Катин Александр Владимирович

Институт развития информационного общества, руководитель дирекции отраслевых программ
РЭУ имени Г. В. Плеханова, старший преподаватель базовой кафедры цифровой экономики ИРИО
Москва, Российская Федерация
alexander.katin@iis.ru

Хохлов Юрий Евгеньевич

Кандидат физико-математических наук, доцент
Институт развития информационного общества, председатель совета директоров
РЭУ имени Г.В. Плеханова, научный руководитель базовой кафедры цифровой экономики ИРИО
Москва, Российская Федерация
yuri.hohlov@iis.ru

Аннотация

Исследованы вопросы мониторинга и оценки уровня развития доверия и безопасности работы с большими данными в России. Разработана концептуальная схема и система показателей для мониторинга мер, предпринимаемых гражданами и организациями по обеспечению доверия и безопасности при работе с большими данными. Для оценки применимости разработанной концептуальной схемы проведена пилотная реализация.

Ключевые слова

большие данные, технологии работы с большими данными, информационная безопасность, доверие

Введение

Вопросы обеспечения информационной безопасности и доверия при использовании цифровых технологий в целом и технологий работы с большими данными в частности являются одними из важнейших факторов, влияющих на процессы цифровой трансформации экономики и общества. Получение ожидаемых социальных и экономических эффектов возможно только в том случае, когда технологии, решения и услуги, основанные на использовании больших данных, безопасны, а пользователи им доверяют.

В Российской Федерации этим вопросам традиционно уделяется повышенное внимание, что нашло свое отражение как в недавно утвержденной Стратегии национальной безопасности Российской Федерации [1], так и в действующей Доктрине информационной безопасности [2], представляющей систему официальных взглядов на обеспечение национальной безопасности Российской Федерации в информационной сфере. Также значимость вопросов обеспечения информационной безопасности отмечается в Стратегии развития информационного общества Российской Федерации до 2030 года [3].

Информационная безопасность Российской Федерации определяется как состояние защищенности личности, общества и государства от внутренних и внешних информационных угроз, при котором обеспечиваются реализация конституционных прав и свобод человека и гражданина, достойные качество и уровень жизни граждан, суверенитет, территориальная целостность и устойчивое социально-экономическое развитие Российской Федерации, оборона и безопасность государства [2]. В рамках настоящей статьи информационная безопасность будет рассматриваться только на уровне процессов, связанных с обеспечением безопасности и доверия

https://doi.org/10.52605/16059921_2021_04_315

при использования цифровых технологий в целом и технологий работы с большими данными в частности. Под доверием будет пониматься готовность пользователей (граждан и организаций) к полноценному применение цифровых технологий и решений с учетом всех возможных угроз и рисков.

Несмотря на наличие стратегических документов, высокий уровень вовлеченности лиц, принимающих решения со стороны органов власти, и выделение значительных финансовых ресурсов [4], в России отсутствует полноценная система мониторинга и оценки уровня доверия и безопасности работы с цифровыми технологиями в целом и с технологиями работы с большими данными в частности. Фактически публичная часть процесса мониторинга ограничивается опросами граждан и организаций в рамках федерального статистического наблюдения [5, 6], связанными лишь с отдельными аспектами информационной безопасности и доверия при использовании цифровых технологий, что не дает возможности проводить даже косвенную оценку уровня доверия и безопасности работы с большими данными.

Информационная безопасность существенно влияет на использование технологий работы с большими данными, поскольку, с одной стороны, может являться барьером (в силу своей сложности и инновационности), а с другой – фактором повышения уровня доверия пользователей (в условиях обеспечения гарантированно высокого уровня безопасности).

Целью данной статьи является разработка компонентов концептуальной схемы мониторинга развития и использования технологий хранения и анализа больших данных в цифровой экономике Российской Федерации (Big Data for Digital Economy, BD4DE) [7], относящихся к оценке уровня доверия и безопасности при работе с большими данными, на основе которой для отдельных показателей мониторинга получены оценки данной сферы в 2020 году.

1 Определение предметной области мониторинга и оценки

Бурное развитие и использование технологий работы с большими данными порождает множество вызовов для обеспечения информационной безопасности. Сверхбольшие объемы данных, их разнообразие и высокая скорость обработки порождают дополнительные риски, зачастую не возникавшие при использовании традиционных ИКТ.

Предметная область мониторинга в рамках настоящий статьи включает в себя два основных компонента: доверие и безопасность при работе с большими данными для граждан и то же самое для организаций.

Уровень доверия и безопасности граждан при работе с большими данными имеет значительный интерес для исследования, поскольку граждане, с одной стороны, являются пользователями технологий, решений и сервисов, основанных на работе с большими данными, а с другой – поставщиками данных. Низкий уровень безопасности данных граждан (прежде всего персональных) порождает множество рисков, связанных с финансовыми и репутационными потерями как для отдельных компаний-владельцев цифровых платформ и сервисов, так и для национальной экономики в целом, поскольку небезопасный оборот персональных данных приводит к повышению уровня преступности в цифровой среде. Слабый уровень доверия граждан к цифровым сервисам может приводить к отказу от их использования, что также является серьезным барьером для развития цифровой экономики.

Доверие и безопасность организаций при работе с большими данными также имеет важное значение, так как технологии, решения и сервисы, основанные на работе с большими данными, все глубже проникают в разнообразные деловые процессы, в то время как традиционные методы защиты информации, применяемые в организациях, не способны отвечать на все вызовы, порождаемые использованием больших данных в повседневной деятельности. Угрозы информационной безопасности, связанные с потерей или несанкционированным раскрытием данных организации, влекут за собой серьезные риски, поскольку, с одной стороны, они способны привести к финансовым и репутационным потерям, а с другой – свести к нулю ожидаемые эффекты от инвестиций в работу с большими данными в организации.

Развитие и использование технологий работы с большими данными в силу многих причин происходит неравномерно и имеет отраслевую специфику, поэтому представляется целесообразным проводить мониторинг и оценку уровня доверия и безопасности работы с большими данными в различных сферах деятельности. Это также важно для проведения межотраслевых сопоставлений. Данные сведения будут полезны как органам власти, отвечающим

за обеспечение информационной безопасности в стране, так и компаниям-поставщикам технологий, решений и сервисов в сфере информационной безопасности при определении направлений дальнейшего развития и использования цифровых технологий.

2 Концептуальная схема и показатели мониторинга доверия и безопасности при работе с большими данными

2.1 Обзор подходов к мониторингу и оценке уровня доверия и безопасности при работе с большими данными

Сегодня отсутствует общепринятая методология мониторинга и оценки уровня доверия и безопасности при работе с большими данными. Существующие подходы к такой деятельности в основном направлены на анализ обеспечения информационной безопасности и доверия применительно к использованию цифровых технологий в целом.

Прежде всего следует отметить деятельность по оценке и сопоставлению уровня развития информационной безопасности в разных странах, осуществляемую в рамках подготовки Глобального индекса кибербезопасности Международного союза электросвязи [8]. Данный индекс отражает уровень обеспечения информационной безопасности по странам, при этом оцениваются следующие виды деятельности по обеспечению кибербезопасности:

- правовые меры;
- технические меры;
- организационные меры;
- меры по развитию потенциала страны в сфере информационной безопасности и доверия;
- меры по организации сотрудничества в этой сфере (международное, межведомственное, межотраслевое).

Индекс публикуется раз в 2 года; в 2020 году Россия заняла 5-е место, набрав 98,06 баллов из 100.

Еще одним подходом к мониторингу и оценке уровня обеспечения кибербезопасности является Национальный индекс кибербезопасности [9] Академии электронного управления, в котором оценивается готовность стран к предотвращению киберугроз и управлению инцидентами, связанными с информационной безопасностью. В основу концептуальной схемы Национального индекса положены угрозы информационной безопасности, реагировать на которые обязана каждая страна: недоступность электронных сервисов; нарушение целостности данных и нарушение конфиденциальности данных. Индекс фокусируется на измеримых аспектах деятельности по обеспечению кибербезопасности среди которых:

- действующее законодательство в сфере кибербезопасности;
- наличие организационных структур, ответственных за обеспечение информационной безопасности;
- механизмы сотрудничества в сфере обеспечение кибербезопасности;
- конкретные результаты деятельности: стратегии и политики, технологические решения, планы развития.

По состоянию на 2020 год в Национальном индексе кибербезопасности Россия находится на 29-м месте из 160.

Всемирным банком в сотрудничестве с Институтом развития информационного общества была разработана методика оценки уровня развития цифровой экономики DECA (Digital Economy Country Assessment) [10], предназначенная для различных стран мира. Одним из факторов, существенно влияющих на развитие цифровой экономики, оцениваемых в рамках указанной методики, выделены доверие и безопасность. Концептуальная схема предметной области данного фактора в DECA включает следующие аспекты:

- государственная политика и регулирование (включает оценку национальной политики в сфере обеспечения информационной безопасности, а также наличие мероприятий, направленных на повышение осведомленности граждан и организаций в сфере информационной безопасности при использовании цифровых технологий);
- организационные меры по обеспечению информационной безопасности (содержит оценку групп реагирования на чрезвычайные ситуации в области информационной

безопасности, а также наличие механизмов государственно-частного партнерства и координации вопросов обеспечения информационной безопасности).

Среди имеющихся подходов к мониторингу и оценке информационной безопасности и доверия следует также отметить деятельность Организации по экономическому сотрудничеству и развитию, которая разработала методологии [11,12] и публикует данные [13] результатов опросов граждан и организаций по вопросам использования информационных технологий, включая аспекты, связанные с оценкой уровня информационной безопасности и доверия.

Несмотря на то, что описанные подходы напрямую не относятся к обеспечению информационной безопасности и доверия при работе с большими данными, основные аспекты приведенных методологий учитываются при формировании концептуальной схемы для целей настоящего исследования.

2.2 Обзор литературы по мониторингу и оценке уровня доверия и безопасности при работе с большими данными

Для разработки концептуальной схемы мониторинга и оценки уровня доверия и безопасности работы с большими данными авторами проведен анализ релевантных научных публикаций, индексируемых в платформе Web of Science (WoS).

С этой целью из «ядерной» коллекции WoS (WoS Core Collection) был выделен массив публикаций в области технологий работы с большими данными на основе следующего поискового образа:

> "big data*" OR bigdata OR "large dataset*" OR "massive data*" OR "data science" OR "data* mining" OR "datamining" OR "text mining" OR "Hadoop*" OR "MapReduce" OR "Map Reduce" OR "unstructured data*" OR "semistructured data*" OR "semi-structured data*" OR "data analytic*" OR "descriptive analytic*" OR "diagnostic analytic*" OR "predictive analytic*" OR "prescriptive analytic*"

Детальное описание процедуры формирования приведенного выше поискового образа описано в статье [7]. За временной период с 2016 по 2020 год с использованием поиска по названию, аннотации, автору и ключевым словам (применено поле «Тема» в инструментарии InCites платформы WoS) было найдено 108 073 публикаций, проиндексированных в WoS.

На следующем этапе из этого массива осуществлялся отбор публикаций, имеющих отношение к доверию и безопасности, путем задания специализированного поискового образа:

> cybersecurity OR "information security" OR trust OR "cybersecurity strategy*" OR "cybersecurity policy*" OR "personal data*" OR "organizational security" OR "technical security" OR "anonymization" OR "privacy by design"

Приведенные ключевые слова, с одной стороны, позволяют получить максимальное покрытие исследуемой области, а с другой – не дают значительного количество «шума», то есть публикаций, слабо связанных с темой исследования (для этого была проведена экспериментальная работа методом экспертных оценок с последующей верификацией по отсеиванию специализированных терминов из первоначального списка, который включал 31 наименование).

Общий и специализированный поисковые запросы были объединены логическим оператором «AND», и в результате после поиска по полю «Тема», было получено 2482 публикации за период с 2016 по 2020 год.

Следующим этапом по обзору публикаций был анализ полученных в результате поиска работ на предмет релевантности исследуемой теме. После анализа наименований, ключевых слов и аннотаций было отобрано 49 публикаций, для которых было принято решение об ознакомлении с полными текстами с целью выделения аспектов, относящихся к мониторингу и оценке доверия и безопасности при работе с большими данными.

Следует отметить, что среди отобранных публикаций встречаются те, что связаны с теоретическими и практическими подходами к использованию технологий работы с большими данными при обеспечении информационной безопасности традиционных информационных систем и цифровой инфраструктуры, однако данное направление не является предметом

настоящего исследования. Изучение полных текстов отобранных публикаций привело к отсеиванию еще 9 работ как нерелевантных.

Оставшиеся 40 публикаций [14-53] можно разделить на две группы:

- публикации, исследующие вопросы доверия и безопасности граждан при работе с большими данными (14 статей);
- работы, исследующие доверие и безопасность при работе с большими данными в организациях (26 статей).

Анализ публикаций из первой группы показал, что основными аспектами, связанными с доверием и безопасностью граждан, являются ограничения и меры защиты, принимаемые гражданами при использовании цифровых сервисов, предоставляемых компаниями, основная деятельность которых связана с обработкой больших данных (социальные сети, маркетплейсы, облачные хранилища, мессенджеры и т.д.). В частности, 10 публикаций в той или иной мере посвящены исследованию ограничений в доступе к функциональности смартфона, принимаемых пользователями при работе с приложениями [14-23]; 13 публикаций исследуют вопросы ограничения пользователей в доступе к своим данным, а также к данным о себе в социальных сетях и облачных хранилищах [14, 16-27]; 9 публикаций относятся к вопросам средств и методов защиты, принимаемых пользователями при работе с цифровыми платформами и сервисами [14, 16-18, 21-24,27].

Анализ публикаций из второй группы, посвященных исследованию доверия и безопасности при работе с большими данными в организациях, позволил выделить следующие аспекты, влияющие на успешность этой деятельности: организационные меры обеспечения информационной безопасности при работе с большими данными (политики, концепции, распорядительные акты, регулирующие информационную безопасность компании в целом, а также безопасность при работе с большими данными в частности) – 13 публикаций [28-40]; технические меры обеспечения информационной безопасности (традиционные [15, 32, 34, 36, 40-42] и специализированные, применяемые при работе с большими данными [13-16, 26, 28, 31-38, 40-42, 45-52]); наличие инцидентов информационной безопасности, связанных с раскрытием, повреждением или уничтожением данных организации – 6 публикаций [15, 27, 31, 35, 40, 53].

Проведенный анализ позволяет сделать вывод, что на сегодня не существует общепризнанных подходов к мониторингу и оценке уровня доверия и безопасности при работе с большими данными для отдельно взятой страны. В то же время явно выделяются два основных направления: исследования вопросов безопасности и доверия граждан при работе с большими данными, а также изучение процессов обеспечения информационной безопасности и доверия в организациях, использующих технологии работы с большими данными. Выделенные направления легли в основу разрабатываемой концептуальной схемы для мониторинга и оценки уровня доверия и безопасности при работе с большими данными.

2.3 Концептуальная схема мониторинга

С учетом проведенного анализа доступных источников, релевантных предмету исследования, была сформирована концептуальная схема мониторинга и оценки уровня доверия и безопасности работы с большими данными в России, которая представлена на рисунке 1.

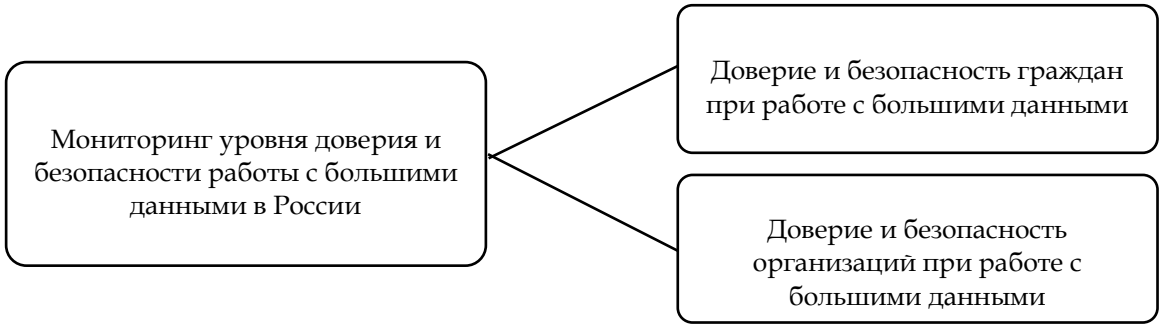

Рисунок 1. Концептуальная схема мониторинга уровня доверия и безопасности работы с большими данными в России

Концептуальная схема включает в себя два компонента: доверие и безопасность граждан при работе с большими данными и доверие и безопасность организаций при работе с большими данными).

2.4 Показатели доверия и безопасности граждан при работе с большими данными

Деятельность пользователей является важным источником больших данных. Множество существующих решений и сервисов основано на обработке и придании добавленной ценности данным, которые фактически производят пользователи в процессе использования приложений и сервисов. Граждане передают компаниям данные о своем местонахождении, выкладывают аудио-, фото- и видеоконтент, осуществляют взаимодействия с другими пользователями через социальные сети и с помощью мессенджеров, хранят свои данные в облачных хранилищах. Важным источником больших данных являются результаты деятельности граждан во Всемирной паутине, которые фиксируются в файлах «куки». Все эти «цифровые следы» наряду с данными, собираемыми поисковыми машинами, являются ценными для маркетинговых компаний, которые на их основе анализируют потребительские предпочтения и формируют персонализированные предложения, а также используют для развития своих продуктов и услуг.

В то же время, полномасштабное развитие технологий и сервисов, основанных на обработке больших массивов пользовательских данных, влечет за собой риски деанонимизации пользователей и некорректного использования их персональных данных, что приводит к потенциально высокому уровню недоверия к использованию такого рода сервисов. В условиях, когда большинство граждан ограничивают доступ к своим данным, развитие сервисов и технологий, основанных на работе с большими данными, крайне затруднительно.

Таким образом, при мониторинге и оценке доверия и безопасности граждан при работе с большими данными предлагается оценить долю граждан, открывающих доступ к своим персональным данным, для чего будут использованы следующие показатели:

- (ДБГ-01-01) Доля граждан, ограничивавших доступ приложениям на своем смартфоне к местоположению, списку контактов, фото/видео контенту, камере, микрофону или отказывавших в таком доступе;
- (ДБГ-01-02) Доля граждан, ограничивавших доступ к сведениям о себе и к своим данным, хранящимся в социальных сетях и (или) в совместно используемых облачных хранилищах.

В рамках оценки уровня доверия и безопасности при работе с большими данными в России видится целесообразным изучить уровень компетенций граждан по защите и/или управлению доступом к своим персональным данным. Чем выше компетенции граждан, тем с большим доверием они будут относиться к использованию сервисов, основанных на больших данных. Мониторинг данного направления позволит оценить, с одной стороны, уровень компетенций граждан по управлению доступом к своим данным, а с другой – уровень доверия пользователей к сервисам, собирающим и использующим их данные.

Для оценки уровня компетенций граждан при управлении доступом к своим данным предлагается использовать следующие показатели:

- (ДБГ-01-03) Доля граждан, изменявших настройки в своем веб-браузере, чтобы предотвратить или ограничить использование файлов «куки»;
- (ДБГ-01-04) Доля граждан, использовавших программное обеспечение, которое ограничивает возможность отслеживания деятельности в интернете.

2.5 Показатели доверия и безопасности при работе с большими данными в организации

Еще одной составляющей данной предметной области мониторинга и оценки является уровень доверия и безопасности при работе с большими данными в организациях. Деятельность по обеспечению информационной безопасности является исключительно важной для всех типов организаций, работающих с большими данными, включая органы государственной власти и местного самоуправления, частные и государственные компании, медицинские и образовательные организации.

Как отмечалось в разделе 1, мониторинг и оценку уровня доверия и безопасности при работе с большими данными в организациях планируется проводить в 10 сферах деятельности:

- система государственного управления;
- здравоохранение;

- образование
- наука;
- промышленность;
- сельское хозяйство;
- строительство;
- развитие городской среды;
- транспорт и логистика;
- энергетическая инфраструктура;
- финансовые услуги.

Основным источником данных для получения значений показателей и анализа ситуации в различных сферах деятельности являются выборочные представительные опросы организаций. Более подробно методология сбора данных и вычисления значений показателей по отдельным предметным областям мониторинга и оценки описана в статье [13].

Возможности технологий работы с большими данными порождают множество рисков информационной безопасности в организациях. Инструменты работы с ними являются зачастую инновационными, что потенциально может привести к невозможности прогнозирования всех потенциальных угроз. Традиционные средства обеспечения информационной безопасности могут не сработать в условиях больших массивов разнообразных данных и значительного количества источников, а также в случае высокой скорости генерации потоков данных. Распределенные хранилища и множество вовлеченного персонала в свою очередь порождают дополнительные риски умышленных или непреднамеренных утечек данных, а также возможность их незапланированного раскрытия.

Исходя из проведенного анализа (см. раздел 2) были выделены три направления мониторинга и оценки уровня безопасности и доверия при работе с большими данными в организации: организационные меры, технические меры, а также частота возникновения в организациях инцидентов, связанных с компрометацией, повреждением или уничтожением данных.

При мониторинге организационных мер обеспечения доверия и безопасности используются два показателя, связанные с наличием в организации политики обеспечения безопасности и доверия – как общей, так и специализированной, направленной на применение технологий работы с большими данными:

- (ДБО-02-01) Доля организаций, имеющих политику обеспечения информационной безопасности;
- (ДБО-02-02) Доля организаций, имеющих специализированную политику обеспечения информационной безопасности для работы с большими данными.

В соответствии с национальным стандартом ГОСТ Р 53114–2008 [54], под политикой (обеспечения) информационной безопасности понимается формальное изложение правил поведения, процедур, практических приемов или руководящих принципов в области информационной безопасности, которыми руководствуется организация в своей деятельности.

Политика обеспечения информационной безопасности при работе с большими данными должна учитывать основные характеристики работы с большими данными, которые порождают дополнительные риски, не учитываемые ранее. Сверхбольшие объемы данных требуют обеспечения информационной безопасности многоуровневых распределенных хранилищ; высокая скорость генерации больших потоков данных требует защиты процессов распределенных вычислений, мониторинга инцидентов в режиме реального времени; однако применение традиционных средств защиты может существенно снизить скорость передачи и обработки данных. Разнообразные форматы структурированных и неструктурированных данных (текст, аудио, видео, изображения) и множество источников их генерации приводят к необходимости применения различных средств защиты, которые должны работать в рамках распределенной инфраструктуры как единое целое.

Политика обеспечения информационной безопасности работы с большими данными может быть утверждена как отдельно, так и быть частью общей политики обеспечения информационной безопасности организации.

Вторым направлением мониторинга и оценки уровня доверия и безопасности при работе с большими данными в организации является оценка интенсивности применения мер обеспечения информационной безопасности в этой организации. В рамках разработанной концептуальной

схемы меры обеспечения информационной безопасности разделяются на базовые (требуемые для обеспечения необходимого уровня информационной защищенности цифровых технологий, используемых организацией) и специализированные (направленные на обеспечение информационной безопасности технологий работы с большими данными). Для этого в концептуальную схему мониторинга BD4DE включены следующие показатели:

- (ДБО-02-03) Доля организаций, применяющих меры обеспечения информационной безопасности при работе с цифровыми технологиями;
- (ДБО-02-04) Доля организаций, применяющих специализированные меры обеспечения информационной безопасности для работы с большими данными.

Под мерой обеспечения информационной безопасности в соответствии с упомянутым выше стандартом [54] понимается совокупность действий, направленных на разработку и/или практическое применение способов и средств обеспечения информационной безопасности. Для оценки показателя ДБО-02-03 была выбрана совокупность мер, используемых организациями, которые отслеживаются Росстатом в рамках федерального статистического наблюдения за использованием информационно-коммуникационных технологий [5].

Еще одним направлением мониторинга в данной предметной области является оценка частоты возникновения инцидентов информационной безопасности, связанных с повреждением или уничтожением данных, а также приведших к утечке конфиденциальных данных. Для оценки этих аспектов в концептуальной схеме BD4DE используются следующие показатели:

- (ДБО-02-05) Доля организаций, сталкивавшихся с инцидентами информационной безопасности, связанными с уничтожением или повреждением данных;
- (ДБО-02-06) Доля организаций, сталкивавшихся с инцидентами информационной безопасности, связанными с раскрытием конфиденциальных данных.

К инцидентам информационной безопасности, приведшим к уничтожению или повреждению данных, могут относиться заражения вредоносным программным обеспечением, несанкционированные проникновения, сбои аппаратного или программного обеспечения.

К инцидентам информационной безопасности, приведшим к утечке конфиденциальных данных, могут относиться несанкционированные проникновения, фарминг (скрытное перенаправление жертвы на ложный IP-адрес), фишинг-атаки (обман пользователей с целью раскрытия конфиденциальных данных), действия собственных сотрудников (преднамеренные или непреднамеренные).

Большое количество инцидентов может влиять на уровень доверия к технологиям работы с большими данными, а также сигнализировать о недостаточном уровне развития организационных и технических мер, принимаемых организацией для обеспечения доверия и безопасности.

3 Методология измерения показателей доверия и безопасности работы с большими данными

Источником данных для расчета значений показателя ДБО-02-03 являются итоги федерального статистического наблюдения за использованием информационно-коммуникационных технологий в организациях.

Для показателей системы мониторинга BD4DE, которые не отслеживаются Федеральной службой государственной статистики, требуется проводить опрос граждан и организаций, для чего разработан соответствующий инструментарий (анкета, паспорта показателей).

Анкета для опроса организаций содержит блок модельных вопросов, позволяющих получить сведения о значениях каждого показателя, предусмотренного для мониторинга и оценки. Анкета содержит определения специализированных терминов, а также пояснения, необходимые для однозначной трактовки задаваемых респондентам вопросов.

Для оценки уровня использования технических мер обеспечения информационной безопасности при работе организаций с цифровыми технологиями (показатель ДБО-02-03) выбраны следующие меры:

- средства строгой аутентификации (например, пароли длиннее 8 символов, сменяемые не реже чем раз в 6 месяцев);
- технические средства аутентификации пользователей (например, токены, USB-ключи или смарт-карты);

- резервное копирование данных на носители, находящиеся физически не на территории организации;
- биометрические средства аутентификации пользователей;
- средства шифрования;
- средства электронной подписи;
- регулярно обновляемые антивирусные программы;
- программные / аппаратные средства, препятствующие несанкционированному доступу вредоносных программ из глобальных информационных / локальных вычислительных сетей (брандмауэр);
- спам-фильтр;
- системы обнаружения вторжения в компьютер или сеть;
- программные средства автоматизации процессов анализа и контроля защищенности компьютерных систем.

Для оценки уровня использования специализированных технических мер защиты, направленных на обеспечение доверия и безопасности при работе с большими данными в организациях (см. показатель ДБО-02-04), выбраны следующие меры:

- средства защиты нереляционных хранилищ данных (например, специализированные алгоритмы криптографического хеширования, такие как SHA-256 или выше, использование протоколов TLS / SSL);
- средства защиты многоуровневых хранилищ данных (например, SUNDR, криптографические облачные хранилища данных);
- криптографические средства обеспечения контроля доступа к данным (например, гомоморфное шифрование, реляционное шифрование, системы, основанные на идентификаторах);
- меры защиты инструментов распределенных вычислений (например, использование протокола Kerberos, технологии мандатного управление доступом);
- средства мониторинга инцидентов и аудита системы обеспечения информационной безопасности в режиме реального времени (например, SIEM решения, средства мониторинга и аудита на всех стадиях жизненного цикла данных);
- средства фильтрации и валидации данных, собираемых из распределенных источников (средства тестирования инфраструктурных ресурсов, использование доверенных сертификатов, использование доверенных устройств и приложений, использование на точках сбора данных средств защиты от вредоносных программ, метрики доверия к узлам и поставщикам данных);
- средства предотвращения деанонимизации пользовательских данных (например, использование технологий дифференциальной конфиденциальности, гомоморфного шифрования);
- другие меры.

4 Результаты пилотной реализации

На первой стадии пилотной реализации было принято решение расширить область мониторинга деятельности организаций показателями, характеризующими уровень развития человеческого капитала организации с точки зрения информационной безопасности. Было предложено дополнить концептуальную схему для данной предметной области показателем, который характеризует наличие у сотрудников организаций навыков по обеспечению доверия и безопасности при работе с большими данными. Другим важным фактором, влияющим на работу с большими данными, является обеспеченность организации специалистами по информационной безопасности, которые способны отвечать на имеющиеся вызовы в сфере информационной безопасности и использовать специализированные инструменты защиты при работе с большими данными.

По результатам первой стадии пилотной реализации были доработаны соответствующие разделы анкет для обследования организаций по вопросам использования технологий работы с большими данными.

На второй стадии пилотной реализации системы мониторинга проводилась экспертная оценка разработанных анкет экспертом данной предметной области (специалистом по

информационной безопасности) и экспертом-социологом (специалистом по разработке анкет и проведению опросов организаций).

Экспертом по информационной безопасности были сформулированы следующие предложения:

- дополнить анкету отдельным вопросом о наличии в организации концепции информационной безопасности, включающей разделы, посвященные безопасности при работе с большими данными;
- дополнить перечень мер обеспечения информационной безопасности при использовании цифровых технологий средствами защиты облачных вычислений;
- дополнить перечень мер обеспечения информационной безопасности при работе с большими данными средствами защиты, поставляемыми сторонними организациями как услуги («Security-as-a-Service» — служба облачных вычислений, обеспечивающая информационную безопасность).

Кроме того, экспертом по информационной безопасности было высказано предположение, что на вопросы анкеты, связанные с обеспеченностью кадрами, большинством респондентов будут даны отрицательные ответы.

Экспертом-социологом были сформулированы следующие предложения: дать пояснения к разделам анкеты о том, какого уровня (в иерархии организации) и какой квалификации специалистов требуется привлечь в качестве респондентов для ответов на вопросы соответствующих разделов; уточнить понятийный аппарат и комментарии по отдельным вопросам.

На последней стадии пилотной реализации полученные замечания и предложения экспертов были учтены в соответствующих разделах финальных версий анкет для опроса организаций.

В настоящее время доступным для расчета показателем, основанным на данных Росстата, является уровень использования технических мер обеспечения информационной безопасности при работе с цифровыми технологиями (показатель ДБО-02-03). Результаты расчета значений показателя по итогам 2019 года [5] для выбранных отраслей представлены ниже на рисунках 2–3.

На рисунке 2 показана доля организаций в различных сферах деятельности, подтвердивших использование хотя бы одной из следующих технических мер при использовании цифровых технологий: средства строгой аутентификации; технические средства аутентификации пользователей; резервное копирование данных на носители, находящиеся физически не на территории организации; биометрические средства аутентификации пользователей.

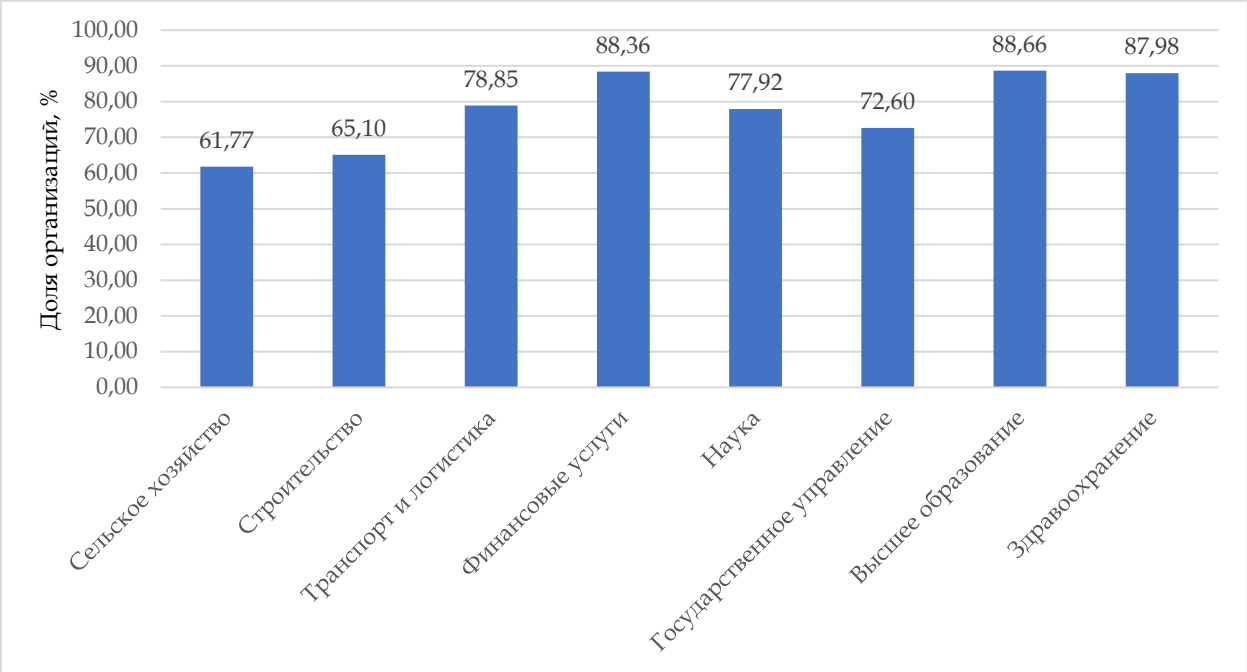

Рисунок 2. Доля российских организаций, использующих средства обеспечения информационной безопасности (в разрезе приоритетных сфер деятельности)

Наиболее популярной мерой из вышеприведенного списка являются технические средства аутентификации пользователей (например, токены, USB-ключи или смарт-карты). Их используют более 50% организаций во всех рассматриваемых сферах деятельности (лидер – высшее образование, 79,43%). Наименее популярной мерой являются биометрические средства аутентификации пользователей. Во всех сферах деятельности, кроме финансовых услуг (24%) уровень использования не превышает 6%.

На рисунке 3 представлены организации в различных сферах деятельности, подтвердившие использование хотя бы одной из следующих технических мер при использовании цифровых технологий:

- средства шифрования;
- средства электронной подписи;
- регулярно обновляемые антивирусные программы;
- программные / аппаратные средства, препятствующие несанкционированному доступу вредоносных программ из глобальных информационных / локальных вычислительных сетей (брандмауэр);
- спам-фильтры;
- системы обнаружения вторжения в компьютер или сеть;
- программные средства автоматизации процессов анализа и контроля защищенности компьютерных систем.

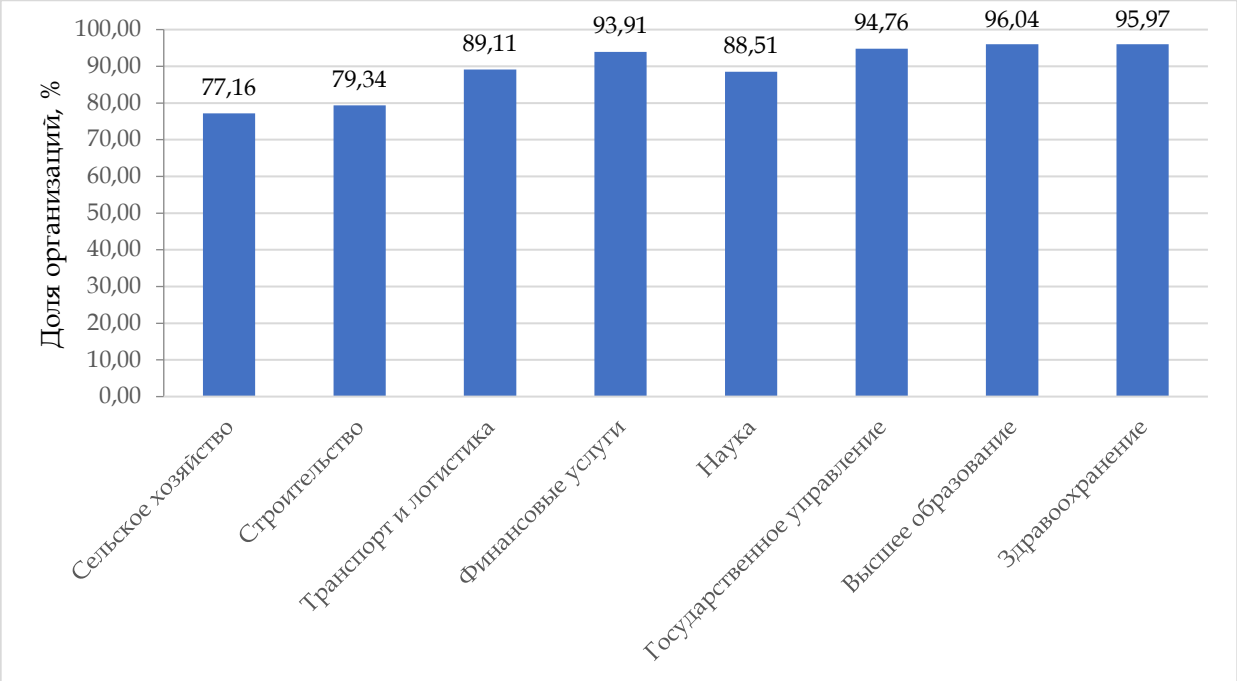

Рисунок 3. Доля российских организаций, использующих средства защиты информации, передаваемой по глобальным сетям (в разрезе приоритетных сфер деятельности)

Наиболее распространенной мерой обеспечения информационной безопасности из приведенного выше списка являются средства электронной подписи, их использование во всех рассматриваемых сферах деятельности находится на уровне 69% и выше. Меньше всего распространены программные средства автоматизации процессов анализа и контроля защищенности компьютерных систем, за исключение финансовой отрасли, в которой данную меру применяют более 60% организаций, и сферы высшего образования (47%); в исследуемых сферах деятельности уровень использования указанного средства обеспечения информационной безопасности не превышает 37%.

Следует отметить, что на графиках не представлены результаты для таких сфер деятельности, как обрабатывающая промышленность, развитие городской среды и энергетическая инфраструктура. Для получения данных по этим отраслям необходимо делать специализированный запрос данных в Главный межрегиональный центр обработки и

распространения статистической информации Росстата России, содержащий сложные группировки кодов ОКВЭД 2, описывающих совокупность видов деятельности, входящих в указанные сферы деятельности, что видится нецелесообразным в рамках пилотной реализации разработанной концептуальной схемы. Сфера деятельности «Образование» ограничена высшим образованием, так как сведения по организациям среднего общего и профессионального образования Росстатом не собираются. Наличие двух графиков для показателя ДБО-02-03 обусловлено спецификой сбора и публикации данных Росстатом, который разделяет технические меры обеспечения информационной безопасности при работе организаций с цифровыми технологиями на два списка.

Понятно, что в полной мере оценить актуальный уровень доверия и безопасности работы с большими данными в России только по указанному показателю не представляется возможным. Данные, необходимые для расчета отобранных показателей, требуется получать в рамках представительного опроса граждан и организаций.

Заключение

В данной работе описана методология мониторинга и оценки уровня доверия и безопасности работы с большими данными в России. Концептуальная схема, предлагаемая для использования в рамках методологии, основывается на изучении релевантных научных публикаций, выведших за последние 5 лет, а также работ ведущих аналитических и исследовательских международных организаций.

Применимость предлагаемой концептуальной схемы и вошедших в нее показателей мониторинга доказана в рамках пилотной реализации с учетом рекомендованных доработок. Следующим этапом апробации предложенного подхода должно стать масштабное полевое исследование, включающее в себя опрос граждан, а также организаций из приоритетных отраслей экономики и социальной сферы с использованием разработанных анкет. В настоящее время данные для расчета значений большей части показателей недоступны.

Дальнейшее развитие настоящего исследования может идти в направлении уточнения концептуальной схемы мониторинга, а также пересмотре перечня показателей, который может потребоваться после сбора всех необходимых данных для проведения оценки, а также получения обратной связи от опрашиваемых организаций и профессионального сообщества.

Благодарности

В работе использованы результаты проекта «Мониторинг и стандартизация развития и использования технологий хранения и анализа больших данных в цифровой экономике Российской Федерации», выполняемого в рамках реализации Программы Центра компетенций Национальной технологической инициативы «Центр хранения и анализа больших данных», поддерживаемого Министерством науки и высшего образования Российской Федерации по Договору МГУ имени М.В.Ломоносова с Фондом поддержки проектов Национальной технологической инициативы от 15.08.2019 № 7/1251/2019.

Работа выполнена при частичной поддержке РФФИ, грант 18-29-03086.

Литература

1. Стратегия национальной безопасности Российской Федерации. Утверждена Указом Президента Российской Федерации от 2 июля 2021 года № 400
2. Доктрина информационной безопасности Российской Федерации. Утверждена Указом Президента Российской Федерации от 5 декабря 2016 г. № 646
3. Стратегия развития информационного общества в Российской Федерации на 2017–2030 годы. Утверждена Указом Президента Российской Федерации от 9 мая 2017 года № 203
4. Национальный проект «Цифровая экономика». Утвержден на заседании президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам 24 декабря 2018 года.
5. Федеральная служба государственной статистики. Итоги федерального статистического наблюдения по ф. № 3-информ «Сведения об использовании информационных и

коммуникационных технологий и производстве вычислительной техники, программного обеспечения и оказании услуг в этих сферах». URL: https://rosstat.gov.ru/folder/14478 (дата обращения 01.06.2021)
6. Федеральная служба государственной статистики. Итоги федерального статистического наблюдения по ф. № 1-ИТ «Анкета выборочного федерального статистического наблюдения по вопросам использования населением информационных технологий и информационно-телекоммуникационных сетей. URL: https://gks.ru/free_doc/new_site/business/it/ikt20/index.html (дата обращения 01.06.2021)
7. Т.В. Ершова, Ю.Е. Хохлов, С.Б. Шапошник. Методология мониторинга развития и использования технологий работы с большими данными // Информационное общество. 2021. № 4–5. С. 2–32. https://doi.org/10.52605/16059921_2021_04_02
8. International Telecommunication Union Development Sector. Global Cybersecurity Index 2020. URL : https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf (дата обращения 01.06.2021)
9. e-Governance Academy Foundation. The National Cyber Security Index. URL: https://ncsi.ega.ee/methodology/ (дата обращения 01.06.2021)
10. Институт развития информационного общества. Анализ текущего состояния развития цифровой экономики в России. М.: Институт развития информационного общества, 2018.— 166 с
11. The OECD Model Survey on ICT Access and Usage by Households and Individuals. 2nd Revision. URL: https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Access-Usage-Households-Individuals.pdf (дата обращения 01.06.2021)
12. The OECD Model Survey on ICT Usage by Businesses. 2nd Revision. URL: https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Usage-Businesses.pdf (дата обращения 01.06.2021)
13. OECD. Data. URL: https://stats.oecd.org/ (дата обращения 01.06.2021)
14. Yuxue Li Lijun Song Yucheng Zeng. Research on information security and privacy protection model based on consumer behavior in big data environment. June 2018. https://doi.org/10.1002/cpe.4881
15. Nils Gruschkay, Vasileios Mavroeidisy, Kamer Vishiy, Meiko Jensen. Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. 2018 IEEE International Conference on Big Data (Big Data). https://doi.org/10.1109/BigData.2018.8622621
16. Luca Bolognini, Camilla Bistolfi, Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation, Computer Law & Security Review: The International Journal of Technology Law and Practice (2016), https://doi.org/10.1016/j.clsr.2016.11.002
17. Grout, V. No More Privacy Any More? Information 2019, 10, 19. https://doi.org/10.3390/info10010019
18. Carole L. Jurkiewicz (2018): Big Data, Big Concerns: Ethics in the Digital Age, Public Integrity, https://doi.org/10.1080/10999922.2018.1448218
19. Alessandro Mantelero, Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework, Computer Law & Security Review: The International Journal of Technology Law and Practice (2017), https://doi.org/10.1016/j.clsr.2017.05.011
20. Martin, K.D., Murphy, P.E. The role of data privacy in marketing. J. of the Acad. Mark. Sci. 45, 135–155 (2017). https://doi.org/10.1007/s11747-016-0495-4
21. Muhammad, S.S., Dey, B.L. & Weerakkody, V. Analysis of Factors that Influence Customers' Willingness to Leave Big Data Digital Footprints on Social Media: A Systematic Review of Literature. Inf Syst Front 20, 559–576 (2018). https://doi.org/10.1007/s10796-017-9802-y
22. N. A. Shozi and J. Mtsweni, "Big data privacy in social media sites," 2017 IST-Africa Week Conference (IST-Africa), 2017, pp. 1-6, https://doi.org/10.23919/ISTAFRICA.2017.8102311.
23. L. Yuqing, "Research on Personal Information Security on Social Network in Big Data Era," 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA), 2017, pp. 676-678, https://doi.org/10.1109/ICSGEA.2017.91.
24. Liu H (2018) Beyond the Scale of Big Data. Front. Big Data 1:1. doi: 10.3389/fdata.2018.00001
25. E. Kosta et al. (Eds.): Privacy and Identity 2018, IFIP AICT 547, pp. 81–94, 2019. https://doi.org/10.1007/978-3-030-16744-8_6

26. Fan, K., Lou, S., Su, R. et al. Secure and private key management scheme in big data networking. Peer-to-Peer Netw. Appl. 11, 992–999 (2018). https://doi.org/10.1007/s12083-017-0579-z
27. SanchengPeng, ShuiYu, PeterMueller. Social networking big data: Opportunities, solutions, and challenges. Future Generation Computer Systems 86 (2018) 1456–1458. doi: https://doi.org/10.1016/j.future.2018.05.040
28. M. Binjubeir, A. A. Ahmed, M. A. B. Ismail, A. S. Sadiq and M. Khurram Khan, "Comprehensive Survey on Big Data Privacy Protection," in IEEE Access, vol. 8, pp. 20067-20079, 2020, https://doi.org/10.1109/ACCESS.2019.2962368.
29. Poltavtseva, M.A. A Consistent Approach to Building Secure Big Data Processing and Storage Systems. Aut. Control Comp. Sci. 53, 914–921 (2019). https://doi.org/10.3103/S0146411619080273
30. Poltavtseva, M.A., Kalinin, M.O. Modeling Big Data Management Systems in Information Security. Aut. Control Comp. Sci. 53, 895–902 (2019). https://doi.org/10.3103/S014641161908025X
31. Hazirah Bee Yusof Ali, Lili Marziana bt Abdullah, Mira Kartiwi, Azlin Nordin. Risk Assessment for Big Data in Cloud: Security, Privacy and Trust. the 2018 Artificial Intelligence and Cloud Computing Conference. https://doi.org/10.1145/3299819.3299841
32. Chang, Y., Government Information Quarterly (2018), https://doi.org/10.1016/j.giq.2018.04.002
33. El Haourani L., Abou El Kalam A., Ait Ouahman A. (2019) Knowledge Based Access Control a Model for Security and Privacy in the Big Data. In: Ben Ahmed M., Boudhir A., Younes A. (eds) Innovations in Smart Cities Applications Edition 2. SCA 2018. Lecture Notes in Intelligent Transportation and Infrastructure. Springer, Cham. https://doi.org/10.1007/978-3-030-11196-0_60
34. N. Miloslavskaya, A. Nikiforov and V. Budzko, Standardization of Ensuring Information Security for Big Data Technologies, 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), 2018, pp. 56-63, https://doi.org/10.1109/W-FiCloud.2018.00015
35. Reza Saneei Moghadam, Ricardo Colomo-Palacios. Information security governance in big data environments: A systematic mapping, Procedia Computer Science, Volume 138, 2018, Pages 401-408, https://doi.org/10.1016/j.procs.2018.10.057
36. Khairulliza Ahmad Salleh, Lech Janczewski, Security Considerations in Big Data Solutions Adoption: Lessons from a Case Study on a Banking Institution, Procedia Computer Science, Volume 164, 2019, Pages 168-176, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.12.169
37. Shetty M.M., Manjaiah D.H., Hemdan E.ED. (2019) Policy-Based Access Control Scheme for Securing Hadoop Ecosystem. In: Balas V., Sharma N., Chakrabarti A. (eds) Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing, vol 839. Springer, Singapore. https://doi.org/10.1007/978-981-13-1274-8_13
38. M. Tang, M. Alazab and Y. Luo, "Big Data for Cybersecurity: Vulnerability Disclosure Trends and Dependencies," in IEEE Transactions on Big Data, vol. 5, no. 3, pp. 317-329, 1 Sept. 2019, https://doi.org/10.1109/TBDATA.2017.2723570
39. F. Xhafa et al. (eds.), Research into Information Security Strategy. Practices for Commercial Banks in Taiwan Recent Developments in Intelligent Systems and Interactive Applications, Advances in Intelligent Systems and Computing 541, https://doi.org/10.1007/978-3-319-49568-2_25
40. A. Yarali, R. Joyce and B. Dixon, "Ethics of Big Data: Privacy, Security and Trust," 2020 Wireless Telecommunications Symposium (WTS), 2020, pp. 1-7, https://doi.org/10.1109/WTS48268.2020.9198734
41. Sahel Alouneh, Ismail Hababeh, Tamer Alajrami. Toward big data analysis to improve enterprise information security. MEDES '18: Proceedings of the 10th International Conference on Management of Digital EcoSystemsSeptember 2018 Pages 106–109 https://doi.org/10.1145/3281375.3281393
42. Macklin, Thomas ; Mathews, Joseph. Big data, little security: Addressing security issues in your platform. Proceedings of the SPIE, Volume 10207, id. 102070G 10 pp. (2017). https://doi.org/10.1117/12.2268002
43. Junjun Guo, Le Wang, Learning to upgrade internet information security and protection strategy in big data era, Computer Communications, Volume 160, 2020, Pages 150-157, https://doi.org/10.1016/j.comcom.2020.05.043
44. Maanak Gupta, Farhan Patwa, Ravi Sandhu . An Attribute-Based Access Control Model for

45. Secure Big Data Processing in Hadoop Ecosystem. the Third ACM Workshop. https://doi.org/10.1145/3180457.3180463
46. Colombo, P., Ferrari, E. Access control technologies for Big Data management systems: literature review and future trends. Cybersecur 2, 3 (2019). https://doi.org/10.1186/s42400-018-0020-9
47. N. Lee and B. Wu, "Privacy Protection Technology and Access Control Mechanism for Medical Big Data," 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2017, pp. 424-429, https://doi.org/10.1109/IIAI-AAI.2017.34
48. D. Mondek, R. B. Blažek and T. Zahradnický, "Security Analytics in the Big Data Era," 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2017, pp. 605-606, https://doi.org/10.1109/QRS-C.2017.136
49. Meisuchi Naisuty, Achmad Nizar Hidayanto, Nabila Clydea Harahap, Ahmad Rosyiq, Agus Suhanto and George Michael Samuel Hartono. Data protection on hadoop distributed file system by using encryption algorithms: a systematic literature review. 2020 J. Phys.: Conf. Ser. 1444 012012. https://doi.org/10.1088/1742-6596/1444/1/012012
50. Salas, J., Domingo-Ferrer, J. Some Basics on Privacy Techniques, Anonymization and their Big Data Challenges. Math.Comput.Sci. 12, 263–274 (2018). https://doi.org/10.1007/s11786-018-0344-6
51. Hai Tao, Md Zakirul Alam Bhuiyan, Md Arafatur Rahman, Guojun Wang, Tian Wang, Md. Manjur Ahmed, Jing Li, Economic perspective analysis of protecting big data security and privacy, Future Generation Computer Systems, Volume 98, 2019, Pages 660-671, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.03.042
52. Linqi Zhou, Weihong Gu, Cheng Huang, Aijun Huang, and Yongbin Bai1. Research on information security in big data era. AIP Conference Proceedings 1967, 020020 (2018). https://doi.org/10.1063/1.5038992
53. Poltavtseva, M.A., Zegzhda, D.P. & Kalinin, M.O. Big Data Management System Security Threat Model. Aut. Control Comp. Sci. 53, 903–913 (2019). https://doi.org/10.3103/S0146411619080261
54. ГОСТ Р 53114–2008. Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения. Утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 18 декабря 2008 г. N 532-ст

TRUST AND SECURITY IN THE SPHERE OF WORKING WITH BIG DATA IN RUSSIA

Katin, Alexander Vladimirovich

Institute of the Information Society, head of Directorate of sectoral programs
Plekhanov Russian University of Economics, IIS-based Digital economy department, senior lecturer
Moscow, Russian Federation
alexander.katin@iis.ru

Hohlov, Yuri Eugenyevich

Candidate of physical and mathematical sciences, associate professor
Institute of the Information Society, chairman of the Board of directors
Plekhanov Russian University of Economics, IIS-Based Digital Economy Department, scientific advisor
Moscow, Russian Federation
yuri.hohlov@iis.ru

Abstract

The issues of monitoring and assessing the level of trust and security in working with big data in Russia have been investigated. A conceptual framework and a system of indicators have been developed for monitoring measures taken by citizens and organizations to ensure trust and security in working with big data. To assess the applicability of the developed framework, a pilot implementation was carried out.

Keywords

big data, big data technologies, information security, trust

References

1. Strategiya nacional`noj bezopasnosti Rossijskoj Federacii. Utverzhdena Ukazom Prezidenta Rossijskoj Federacii ot 2 iyulya 2021 goda № 400
2. Doktrina informacionnoj bezopasnosti Rossijskoj Federacii. Utverzhdena Ukazom Prezidenta Rossijskoj Federacii ot 5 dekabrya 2016 g. № 646
3. Strategiya razvitiya informacionnogo obshhestva v Rossijskoj Federacii na 2017- 2030 gody`. Utverzhdena Ukazom Prezidenta Rossijskoj Federacii ot 9 maya 2017 goda № 203
4. Nacional`ny`j proekt «Cifrovaya e`konomika». Utverzhden na zasedanii prezidiuma Soveta pri Prezidente Rossijskoj Federacii po strategicheskomu razvitiyu i nacional`ny`m proektam 24 dekabrya 2018 goda.
5. Federal`naya sluzhba gosudarstvennoj statistiki. Itogi federal`nogo statisticheskogo nablyudeniya po f. № 3-inform «Svedeniya ob ispol`zovanii informacionny`x i kommunikacionny`x texnologij i proizvodstve vy`chislitel`noj texniki, programmnogo obespecheniya i okazanii uslug v e`tix sferax». URL: https://rosstat.gov.ru/folder/14478 (accessed on 01.06.2021).
6. Federal`naya sluzhba gosudarstvennoj statistiki. Itogi federal`nogo statisticheskogo nablyudeniya po f. № 1-IT «Anketa vy`borochnogo federal`nogo statisticheskogo nablyudeniya po voprosam ispol`zovaniya naseleniem informacionny`x texnologij i informacionno-telekommunikacionny`x setej. URL: https://gks.ru/free_doc/new_site/business/it/ikt20/index.html (accessed on 01.06.2021)
7. T.V. Ershova, Yu.E. Hohlov, S.B. Shaposhnik. Metodologiya monitoringa razvitiya i ispol`zovaniya texnologij raboty` s bol`shimi danny`mi // Informacionnoe obshhestvo. 2021. № 4–5. S. 2–32. https://doi.org/10.52605/16059921_2021_04_02
8. International Telecommunication Union Development Sector. Global Cybersecurity Index 2020. URL: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf (accessed on 01.06.2021).
9. e-Governance Academy Foundation. The National Cyber Security Index. URL: https://ncsi.ega.ee/methodology/ (accessed on 01.06.2021)/

10. Institut razvitiya informacionnogo obshhestva. Analiz tekushhego sostoyaniya razvitiya cifrovoj e`konomiki v Rossii. M.: Institut razvitiya informacionnogo obshhestva, 2018.— 166 s.
11. The OECD Model Survey on ICT Access and Usage by Households and Individuals. 2nd Revision. URL: https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Access-Usage-Households-Individuals.pdf (accessed on 01.06.2021)/
12. The OECD Model Survey on ICT Usage by Businesses. 2nd Revision. URL: https://www.oecd.org/sti/ieconomy/ICT-Model-Survey-Usage-Businesses.pdf (accessed on 01.06.2021).
13. OECD. Data. URL: https://stats.oecd.org/ (accessed on 01.06.2021)/
14. Yuxue Li Lijun Song Yucheng Zeng. Research on information security and privacy protection model based on consumer behavior in big data environment. June 2018. https://doi.org/10.1002/cpe.4881
15. Nils Gruschkay, Vasileios Mavroeidisy, Kamer Vishiy, Meiko Jensen. Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. 2018 IEEE International Conference on Big Data (Big Data). https://doi.org/10.1109/BigData.2018.8622621
16. Luca Bolognini, Camilla Bistolfi, Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation, Computer Law & Security Review: The International Journal of Technology Law and Practice (2016), https://doi.org/10.1016/j.clsr.2016.11.002
17. Grout, V. No More Privacy Any More? Information 2019, 10, 19. https://doi.org/10.3390/info10010019
18. Carole L. Jurkiewicz (2018): Big Data, Big Concerns: Ethics in the Digital Age, Public Integrity, https://doi.org/10.1080/10999922.2018.1448218
19. Alessandro Mantelero, Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework, Computer Law & Security Review: The International Journal of Technology Law and Practice (2017), https://doi.org/10.1016/j.clsr.2017.05.011
20. Martin, K.D., Murphy, P.E. The role of data privacy in marketing. J. of the Acad. Mark. Sci. 45, 135–155 (2017). https://doi.org/10.1007/s11747-016-0495-4
21. Muhammad, S.S., Dey, B.L. & Weerakkody, V. Analysis of Factors that Influence Customers' Willingness to Leave Big Data Digital Footprints on Social Media: A Systematic Review of Literature. Inf Syst Front 20, 559–576 (2018). https://doi.org/10.1007/s10796-017-9802-y
22. N. A. Shozi and J. Mtsweni, Big data privacy in social media sites, 2017 IST-Africa Week Conference (IST-Africa), 2017, pp. 1-6, https://doi.org/10.23919/ISTAFRICA.2017.8102311.
23. L. Yuqing, Research on Personal Information Security on Social Network in Big Data Era, 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA), 2017, pp. 676-678, https://doi.org/10.1109/ICSGEA.2017.91
24. Liu H (2018) Beyond the Scale of Big Data. Front. Big Data 1:1. doi: 10.3389/fdata.2018.00001
25. E. Kosta et al. (Eds.): Privacy and Identity 2018, IFIP AICT 547, pp. 81–94, 2019. https://doi.org/10.1007/978-3-030-16744-8_6
26. Fan, K., Lou, S., Su, R. et al. Secure and private key management scheme in big data networking. Peer-to-Peer Netw. Appl. 11, 992–999 (2018). https://doi.org/10.1007/s12083-017-0579-z
27. SanchengPeng, ShuiYu, PeterMueller. Social networking big data: Opportunities, solutions, and challenges. Future Generation Computer Systems 86 (2018) 1456–1458. doi: https://doi.org/10.1016/j.future.2018.05.040
28. M. Binjubeir, A. A. Ahmed, M. A. B. Ismail, A. S. Sadiq and M. Khurram Khan, Comprehensive Survey on Big Data Privacy Protection, in IEEE Access, vol. 8, pp. 20067-20079, 2020, https://doi.org/10.1109/ACCESS.2019.2962368
29. Poltavtseva, M.A. A Consistent Approach to Building Secure Big Data Processing and Storage Systems. Aut. Control Comp. Sci. 53, 914–921 (2019). https://doi.org/10.3103/S0146411619080273
30. Poltavtseva, M.A., Kalinin, M.O. Modeling Big Data Management Systems in Information Security. Aut. Control Comp. Sci. 53, 895–902 (2019). https://doi.org/10.3103/S014641161908025X
31. Hazirah Bee Yusof Ali, Lili Marziana bt Abdullah, Mira Kartiwi, Azlin Nordin. Risk Assessment for Big Data in Cloud: Security, Privacy and Trust. the 2018 Artificial Intelligence and Cloud Computing Conference. https://doi.org/10.1145/3299819.3299841

32. Chang, Y., Government Information Quarterly (2018), https://doi.org/10.1016/j.giq.2018.04.002
33. El Haourani L., Abou El Kalam A., Ait Ouahman A. (2019) Knowledge Based Access Control a Model for Security and Privacy in the Big Data. In: Ben Ahmed M., Boudhir A., Younes A. (eds) Innovations in Smart Cities Applications Edition 2. SCA 2018. Lecture Notes in Intelligent Transportation and Infrastructure. Springer, Cham. https://doi.org/10.1007/978-3-030-11196-0_60
34. N. Miloslavskaya, A. Nikiforov and V. Budzko, Standardization of Ensuring Information Security for Big Data Technologies, 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), 2018, pp. 56-63, https://doi.org/10.1109/W-FiCloud.2018.00015
35. Reza Saneei Moghadam, Ricardo Colomo-Palacios. Information security governance in big data environments: A systematic mapping, Procedia Computer Science, Volume 138, 2018, Pages 401-408, https://doi.org/10.1016/j.procs.2018.10.057
36. Khairulliza Ahmad Salleh, Lech Janczewski, Security Considerations in Big Data Solutions Adoption: Lessons from a Case Study on a Banking Institution, Procedia Computer Science, Volume 164, 2019, Pages 168-176, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.12.169
37. Shetty M.M., Manjaiah D.H., Hemdan E.ED. (2019) Policy-Based Access Control Scheme for Securing Hadoop Ecosystem. In: Balas V., Sharma N., Chakrabarti A. (eds) Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing, vol 839. Springer, Singapore. https://doi.org/10.1007/978-981-13-1274-8_13
38. M. Tang, M. Alazab and Y. Luo, Big Data for Cybersecurity: Vulnerability Disclosure Trends and Dependencies, in IEEE Transactions on Big Data, vol. 5, no. 3, pp. 317-329, 1 Sept. 2019, https://doi.org/10.1109/TBDATA.2017.2723570
39. F. Xhafa et al. (eds.), Research into Information Security Strategy. Practices for Commercial Banks in Taiwan Recent Developments in Intelligent Systems and Interactive Applications, Advances in Intelligent Systems and Computing 541, https://doi.org/10.1007/978-3-319-49568-2_25
40. A. Yarali, R. Joyce and B. Dixon, Ethics of Big Data: Privacy, Security and Trust, 2020 Wireless Telecommunications Symposium (WTS), 2020, pp. 1-7, https://doi.org/10.1109/WTS48268.2020.9198734
41. Sahel Alouneh, Ismail Hababeh, Tamer Alajrami. Toward big data analysis to improve enterprise information security. MEDES '18: Proceedings of the 10th International Conference on Management of Digital EcoSystemsSeptember 2018, Pages 106–109, https://doi.org/10.1145/3281375.3281393
42. Macklin, Thomas ; Mathews, Joseph. Big data, little security: Addressing security issues in your platform. Proceedings of the SPIE, Volume 10207, id. 102070G 10 pp. (2017). https://doi.org/10.1117/12.2268002
43. Junjun Guo, Le Wang, Learning to upgrade internet information security and protection strategy in big data era, Computer Communications, Volume 160, 2020, Pages 150-157, https://doi.org/10.1016/j.comcom.2020.05.043
44. Maanak Gupta, Farhan Patwa, Ravi Sandhu . An Attribute-Based Access Control Model for
45. Secure Big Data Processing in Hadoop Ecosystem. the Third ACM Workshop. https://doi.org/10.1145/3180457.3180463
46. Colombo, P., Ferrari, E. Access control technologies for Big Data management systems: literature review and future trends. Cybersecur 2, 3 (2019). https://doi.org/10.1186/s42400-018-0020-9
47. N. Lee and B. Wu, Privacy Protection Technology and Access Control Mechanism for Medical Big Data, 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2017, pp. 424-429, https://doi.org/10.1109/IIAI-AAI.2017.34
48. D. Mondek, R. B. Blažek and T. Zahradnický, "Security Analytics in the Big Data Era," 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2017, pp. 605-606, https://doi.org/10.1109/QRS-C.2017.136
49. Meisuchi Naisuty, Achmad Nizar Hidayanto, Nabila Clydea Harahap, Ahmad Rosyiq, Agus Suhanto and George Michael Samuel Hartono. Data protection on hadoop distributed file system by using encryption algorithms: a systematic literature review. 2020 J. Phys.: Conf. Ser. 1444 012012. https://doi.org/10.1088/1742-6596/1444/1/012012
50. Salas, J., Domingo-Ferrer, J. Some Basics on Privacy Techniques, Anonymization and their Big Data Challenges. Math.Comput.Sci. 12, 263–274 (2018). https://doi.org/10.1007/s11786-018-0344-6

51. Hai Tao, Md Zakirul Alam Bhuiyan, Md Arafatur Rahman, Guojun Wang, Tian Wang, Md. Manjur Ahmed, Jing Li, Economic perspective analysis of protecting big data security and privacy, Future Generation Computer Systems, Volume 98, 2019, Pages 660-671, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.03.042
52. Linqi Zhou, Weihong Gu, Cheng Huang, Aijun Huang, and Yongbin Bai1. Research on information security in big data era. AIP Conference Proceedings 1967, 020020 (2018). https://doi.org/10.1063/1.5038992
53. Poltavtseva, M.A., Zegzhda, D.P. & Kalinin, M.O. Big Data Management System Security Threat Model. Aut. Control Comp. Sci. 53, 903–913 (2019). https://doi.org/10.3103/S0146411619080261
54. GOST R 53114-2008. Zashhita informacii. Obespechenie informacionnoj bezopasnosti v organizacii. Osnovny`e terminy` i opredeleniya. Utverzhden i vveden v dejstvie Prikazom Federal`nogo agentstva po texnicheskomu regulirovaniyu i metrologii ot 18 dekabrya 2008 g. N 532-st