

Информация для цитирования:

Мартынов А. В., Ширеева Е. В. Правовые и практические вопросы использования синтетических данных для целей государственного управления в России // Вестник Пермского университета. Юридические науки. 2025. Вып. 4(70). С. 526–538. DOI: 10.17072/1995-4190-2025-70-526-538.

Martynov A. V., Shireeva E. V. *Pravovye i prakticheskie voprosy ispol'zovaniya sinteticheskikh dannykh dlya tseley gosudarstvennogo upravleniya v Rossii* [Legal and Practical Issues of Synthetic Data Use for Public Administration Purposes in Russia]. *Vestnik Permskogo universiteta. Juridicheskie nauki* – Perm University Herald. Juridical Sciences. 2025. Issue 4(70). Pp. 526–538. (In Russ.). DOI: 10.17072/1995-4190-2025-70-526-538.

УДК 342

DOI: 10.17072/1995-4190-2025-70-526-538

Правовые и практические вопросы использования синтетических данных для целей государственного управления в России

*Исследование выполнено за счет гранта Российского научного фонда № 25-28-00491,
<https://rscf.ru/project/25-28-00491/>*

А. В. Мартынов

Национальный исследовательский
Нижегородский государственный университет
имени Н. И. Лобачевского
E-mail: docpred@yandex.ru

Е. В. Ширеева

Национальный исследовательский
Нижегородский государственный университет
имени Н. И. Лобачевского
E-mail: shireevaekaterina@yandex.ru

Статья поступила в редакцию 10.07.2025

Введение: данные становятся основным стратегическим ресурсом, без которого невозможен качественный переход к развитому генеративному искусственному интеллекту и совершенствование методов и подходов к работе с большими данными. Однако обеспечение доступа к большим объемам данных широкого круга пользователей сталкивается с рядом существенных и очень устойчивых правовых барьеров. **Цель** настоящего исследования – определить практическое значение публичных цифровых озер для создания синтетических данных, а также рассмотреть перспективы их использования для целей публичного управления. **Методы:** исследование базируется на совокупности общенаучных (анализ, синтез, индукция, дедукция), частнонаучных (статистический, социологический) и специальных методов познания (формально-юридический, сравнительно-правовой, метод правового моделирования). **Результаты и выводы:** ключевым вопросом становится определение прав в отношении данных (установление обладателя данных) и наличия в массивах данных чувствительной информации, общий доступ к которой, в силу ряда причин, невозможен и несет в себе существенные риски для государства, общества или личности. В рамках этой тенденции оптимальным решением видится формирование крупных массивов данных (озер данных), управляемых государством, а также распространение практики использования синтетических данных. Россия на пути к экономике данных активно формирует публичные озера данных, которые в потенциале могут стать основой для создания на их основе синтетических данных. С этой целью был проанализирован российский и зарубежный опыт правового регулирования в указанной сфере, выявлены и сформулированы практические и правовые проблемы применения цифровых озер для синтетических данных в целях публичного управления.

© Мартынов А. В., Ширеева Е. В., 2025



Данная работа распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>

Ключевые слова: синтетические данные; озеро данных; государственное управление; органы исполнительной власти; искусственный интеллект; правовое регулирование; генеративный искусственный интеллект; машинное обучение

Legal and Practical Issues of Synthetic Data Use for Public Administration Purposes in Russia

The study was supported by a grant from the Russian Science Foundation, project No. 25-28-00491, <https://rscf.ru/project/25-28-00491/>

A. V. Martynov

Lobachevsky State University of Nizhny Novgorod
E-mail: docpred@yandex.ru

E. V. Shireeva

Lobachevsky State University of Nizhny Novgorod
E-mail: shireevaekaterina@yandex.ru

Received 10 Jul 2025

Introduction: data are becoming the main strategic source, without which it is impossible to make a qualitative transition to advanced generative artificial intelligence and improve methods and approaches to big data processing. However, there are some significant and deep-rooted legal barriers on the way to providing access to large amounts of data for a wide range of users. **Purpose:** the study aims to substantiate the practical importance of public data lakes for creating synthetic data as well as to explore the prospects for their use for public administration purposes. **Methods:** the research is based on a combination of general scientific methods (analysis, synthesis, induction, deduction), special scientific methods (statistical, sociological), and specialized methods of cognition (formal legal, comparative legal methods, legal modeling). **Results and conclusions:** the key issues are the rights in relation to data (who owns the data) and the presence in datasets of sensitive information that, for a number of reasons, must not be openly shared and carries significant risks for the state, society, or the individual. What seems to be the most optimal solution is the formation of large arrays of data (data lakes) managed by the state as well as the spread of the practice of using synthetic data. On its way to a data economy, Russia is actively developing the practice of creation of public data lakes, which can potentially become the source for creating synthetic data. The paper analyzes Russian and foreign experience of legal regulation in this area, identifying practical and legal issues of the use of synthetic data for public administration purposes.

Keywords: synthetic data; data lake; public administration; executive authorities; artificial intelligence; legal regulation; generative artificial intelligence; machine learning

Введение

Развитие технологии искусственного интеллекта (ИИ) высокого уровня порождает необходимость получения и обработки огромного массива информации. Она используется для обучения искусственного интеллекта как в ручном, так и в автономном режиме. При этом образуемые нейросети функционируют не только на основе существующей информации, полученной разными способами и через различные источники, но и посредством создания новых данных, которые не имеют естественного (исходящего от человека) происхождения.

На современном этапе развития цифровых технологий именно генеративный искусственный интеллект считается искусственным интеллектом высокого уровня, то есть способным преобразовывать существующую информацию и создавать абсолютно новые данные. Эти данные (информация) наиболее приближены к реальным данным (информации),

создаваемым человеком (литературное произведение, служебный документ, картина, фотография, изображение и т. п.). При этом обеспечивается необходимый уровень достоверности новых данных, что свидетельствует об их приемлемости и допустимости в различных общественных процессах.

Речь идет об информации, созданной или преобразованной генеративным искусственным интеллектом. Такая информация (данные, сведения, цифровые объекты) получила наименование синтетических данных. В самом упрощенном виде под синтетическими данными обычно понимают информацию, сгенерированную компьютерами и по своим статистическим свойствам напоминающую реальные данные. Она создается с помощью алгоритмов без прямой связи с реальными сведениями [15].

Илон Маск в интервью газете «Гардиан» заявил, что у компаний, занимающихся искусственным интеллектом, закончились данные для обучения их моделей и что они «исчерпали» все человеческие

знания. Он предположил, что технологическим компаниям придется обратиться к синтетическим данным – или материалам, созданным моделями искусственного интеллекта, – чтобы создавать и настраивать новые системы [14]. Эндрю Дункан, директор по науке и инновациям (фундаментальные исследования в области DS и AI) в британском Институте Алана Тьюринга, подтвердил, что комментарий Маска согласуется с недавней научной работой, в которой говорится, что общедоступные данные для моделей ИИ могут закончиться уже в 2026 году [17]. Он добавил, что чрезмерная зависимость от синтетических данных способна привести к «краху модели» – так называют ухудшение качества результатов работы моделей [14]. Вместе с тем уже сейчас могут быть применены весьма эффективные решения для преодоления этой проблемы. Речь идет о цифровых озерах данных (Data Lake) как источнике данных для обучения генеративного искусственного интеллекта.

Озеро данных – это, как правило, единое хранилище данных, включающее необработанные копии данных исходной системы, данные датчиков, социальные данные и т. д., а также преобразованные данные, используемые для составления отчетов, визуализации, расширенной аналитики и машинного обучения и т. п.

Методологическая основа исследования

Исследование базируется на совокупности общенаучных (анализ, синтез, индукция, дедукция), частнонаучных (статистический, социологический) и специальных методов познания (формально-юридический, сравнительно-правовой, метод правового моделирования). Поскольку тема исследования имеет правовую специфику, остановимся на характеристике лишь специальных методов познания, которые присущи исключительно юриспруденции. Формально-юридический метод применялся для уточнения содержания норм права: он позволил их истолковать, определить структуру и терминологию законов. Кроме того, с его помощью были выявлены теоретические и практические проблемы применения синтетических данных и цифровых озер в деятельности публичных органов. Сравнительно-правовой метод использовался для соотнесения правового регулирования и практики применения синтетических данных и цифровых озер в публичном секторе в России, США, Европе и Китае. Он позволил выделить критерии для сравнения правового регулирования, выявить специфические и общие черты в практике применения публичными органами цифровых озер и синтетических данных. Поскольку в исследовании определены проблемы научного и правового характера, применение метода правового моделирования позволило сформулировать решение этих проблем, а также выводы по научному исследованию в целом.

Теоретическая основа исследования

Следует отметить, что исследования о правовом регулировании применения цифровых озер и синтетических данных в публичном секторе фактически отсутствуют. Это связано с тем, что данные технологии возникли в публичном секторе большинства государств лишь в 2022–2024 годах. Есть лишь исследования по некоторым аспектам применения цифровых озер и синтетических данных.

Теоретические основы цифровых озер описаны Джеймсом Диксоном как альтернатива «хранилищу данных» [8]. Структура цифровых озер и их технические принципы работы в публичном секторе рассмотрены в исследованиях Билла Инмона [10]. Роль цифровых озер в деятельности органов государственной власти изучена Робом Китчином [12], а также Инес Мергель, Ноэллой Эдельманн и Натали Хауг [13]. Имеются, кроме того, исследования, посвященные опыту применения цифровых озер в конкретных государствах [1; 6; 9; 18]. Вопросы неприкосновенности частной жизни и риски нарушения конфиденциальности тоже становились предметом исследования [7; 9; 18].

Настоящее исследование освещает не только вышеназванные аспекты. Его спецификой является рассмотрение вопроса через призму проблематики публичного управления с учетом правового опыта и практики применения цифровых озер и синтетических данных в Российской Федерации.

Международный опыт. Регулирование синтетических данных и озер данных. Искусственный интеллект

1. США

Использование озер данных государственными структурами в США имеет довольно давнюю историю. Реализованные инициативы и решения были направлены на создание или улучшение использования правительственными агентствами больших массивов данных.

С учетом этого появляется все больше проектов, направленных на интеграцию разрозненных хранилищ данных в единые правительственные озера данных, а также привлечение к этому частных компаний для выполнения подрядных работ. Примером могут служить инициативы Министерства обороны США (US Department of Defense) и Бюро переписи населения США (US Census Bureau), которые в 2024 году объявили о серьезных планах интеграции структурированных, полуструктурированных и «сырых» озер данных [6]. Возрастает у федеральных агентств США интерес и к созданию публичных, то есть открытых, озер данных, которые являются своего рода продолжением уже давно осознанной тенденции обеспечения открытости данных. В частности, это подтверждает создание Public-Use Data Lake, инициированное Министерством труда США (US Department of Labour).

Современные потребности в применении технологий генеративного искусственного интеллекта в государственном управлении как на федеральном уровне, так и на уровне штатов диктуют необходимость крайне разборчиво подходить к выбору этих технологий по причине высоких рисков их влияния одновременно на интересы государственного управления и обеспечение прав граждан. Сегодня уже очевидно, что применение реальных данных грозит раскрытием чувствительной информации и незаконным ее использованием с учетом существующих ограничений в доступе [18]. Наиболее вероятным выходом из ситуации представляется использование для тестирования рассматриваемых к применению технологий генеративного ИИ искусственных данных и озер данных.

В частности, уже есть конкретные примеры сотрудничества локальных правительств с университетскими исследовательскими центрами и частными компаниями. Так, Центр системных данных Мэриленда (Maryland Longitudinal Data System Center) сотрудничает с несколькими исследовательскими центрами для формирования на основе образовательных данных по школьникам и студентам округа искусственных данных для последующего их использования в целях улучшения публичных услуг в этой сфере. Другим примером может служить сотрудничество Института городского развития (Urban Institute) с Департаментом социальных служб округа Аллегейни (Allegheny County Department of Human Services) и Региональным центром данных Западной Пенсильвании (Western Pennsylvania Regional Data Center) в рамках пилотного проекта по использованию синтетических данных для улучшения социальных слуг и программ [16].

На федеральном уровне также отмечен интерес к проблематике использования искусственных данных в различных целях. В январе 2024 года Федеральный совет руководителей служб обработки данных (Federal Chief Data Officers Council) опубликовал запрос на сотрудничество в проведении исследований в области искусственных данных, отметив, что современное понимание этого явления и подходы к нему не отвечают актуальным потребностям государства и бизнеса в развитии и регулировании искусственных данных, имеющих необходимый потенциал для экономики.

2. Европа

В Европейском союзе присутствует выраженный интерес к правовому регулированию вопросов использования технологий искусственного интеллекта и оборота данных в сфере публичного управления. В этой связи в дополнение к уже широко известному Общему регламенту по защите данных (General Data Protection Regulation)¹ были приняты Закон об искусственном интеллекте (Artificial Intelligence Act)² и Закон о данных (Data Act)³, которые в совокупности определяют общую стратегию для стран Европы как в части управления данными, так и в части использования технологий ИИ.

Так, Закон об искусственном интеллекте содержит упоминание о том, что не существует непреодолимых препятствий для развития, внедрения и использования технологий искусственного интеллекта в сфере государственного управления и предоставления публичных услуг. Однако при этом должны быть соблюдены все необходимые фундаментальные права и свободы личности. В частности, Европейский надзорный орган по защите данных (European Data Protection Supervisor) в 2023 году опубликовал первый набор рекомендаций (Guidelines) по использованию технологий генеративного искусственного интеллекта для органов и учреждений Евросоюза⁴. Несмотря на то что документ в принципе имеет рекомендательный характер, он содержит вполне четкие и конкретные правила, которые позволяют провести оценку потенциальных рисков применения указанных технологий.

Персонал органов Евросоюза должен придерживаться как минимум пяти важных принципов при использовании генеративного искусственного интеллекта:

- служащий не вправе разглашать и передавать для использования генеративным искусственным интеллектом информацию, которая не имеет статуса общедоступной или публичной;
- следует критично относиться к полученному результату в части его точности и релевантности, а также наличия в нем противоречий или ложной информации;
- следует критически подходить к использованию публичных генеративных технологий искусственного

¹ General Data Protection Regulation. URL: <https://eur-lex.europa.eu/EN/legal-content/summary/general-data-protection-regulation-gdpr.html>.

² Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>.

³ Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act). URL: <https://eur-lex.europa.eu/eli/reg/2023/2854>.

⁴ First EDPS Orientations for ensuring data protection compliance when using Generative AI systems, European Data Protection Supervisor. URL: https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf.

интеллекта во избежание нарушения прав на интеллектуальную собственность;

– нельзя использовать генеративный искусственный интеллект для создания юридически значимых или обязывающих документов;

– нельзя использовать генеративный искусственный интеллект для критических и «чувствительных» процессов.

Сам по себе Закон об искусственном интеллекте содержит ряд важных правил использования искусственного интеллекта для целей публичного управления, а именно связанных с применением особой маркировки как для результатов, полученных с помощью искусственного интеллекта, так и для случаев взаимодействия гражданина и искусственного интеллекта при оказании публичных услуг.

Закон о данных также внес серьезные коррективы в регулирование оборота данных в рамках Европейского союза. С учетом целей и задач настоящего исследования необходимо отметить то, что данный документ предоставил органам власти право получить доступ к данным, используемым и накапливаемым субъектами частного сектора, при определенных условиях чрезвычайных ситуаций.

Еще одним документом, который активно влияет на создание рынка данных в Европейском союзе, является Директива об открытых данных (Open Data Directive)¹, которая четко закрепляет необходимость раскрытия информации, собираемой органами власти по шести тематическим областям (геопространственная; наблюдение за Землей и окружающей средой; метеорологическая; статистическая; компании и собственность компаний; мобильность). Наборы данных должны предоставляться бесплатно в машиночитаемых форматах и через API для облегчения их эффективного повторного использования. В целом реализация Директивы привела к появлению и развитию European Data Portal – общеевропейской многоязычной платформы для поиска информации государственного сектора, открытой для повторного использования.

В рамках рассматриваемой стратегии, в части построения рынка данных в Европейском союзе, синтетические данные все чаще рассматриваются как аналог анонимизированных данных для снижения риска нанесения ущерба правам и интересам физических лиц при обработке персональных данных [9], о чем более подробно изложено в разделе «Неприкосновенность частной жизни / риски нарушения конфиденциальности».

3. Kumai

Китай известен как страна, которая в первую очередь опирается на концепцию суверенитета данных и рассматривает данные как стратегический ресурс своей национальной экономики данных. Правовое регулирование национальной системы данных включает в себя три основных документа: Закон о кибербезопасности (Cybersecurity Law)², Закон о защите данных (Data Security Law)³ и Закон о защите личной информации (Personal Information Protection Law)⁴.

С точки зрения регулирования вопросов оборота данных наибольший интерес представляют ключевые положения Закона о защите данных. Обратимся к наиболее важным среди них:

– *категоризация данных*. Данные классифицируются исходя из их значения для национальной безопасности, публичного интереса и развития экономики, что позволяет правительству использовать различные охранительные режимы в зависимости от их чувствительности и потенциального влияния;

– *доступ государства к данным*. В рамках сохранения контроля над данными резидентов страны правительство обладает существенными полномочиями по доступу к этим данным, в том числе и обрабатываемым в частном секторе;

– *локализация данных и ограничение трансграничной передачи*. Вся чувствительная и критическая информация должны генерироваться и обрабатываться под юрисдикцией государства. Трансграничная передача допускается только при условии получения на то специального разрешения компетентного органа;

– *эффективное использование данных*. В противовес ограничениям внедряются инструменты поддержки развития технологий по обработке данных при условии, что они соответствуют установленным правилам.

Безусловно, данный обзор не дает общего представления о возможности использования синтетических данных, как и об их регулировании и применении в целях государственного управления, однако позволяет сформировать общее понимание существующих тенденций в правовом регулировании и их направленности в данной сфере.

Стоит отметить и то, что такое направление государственной политики демонстрирует создание единого подхода к обработке правительственных данных и их интеграции в единые озера данных. В частности, в 2022 году было анонсировано создание Интегрированной государственной системы больших данных (Integrated Government Big Data

¹ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.172.01.0056.01.ENG.

² Cybersecurity Law of the People's Republic of China (unofficial translation). URL: https://www.dataguidance.com/sites/default/files/en_cybersecurity_law_of_the_peoples_republic_of_china_1.pdf.

³ Data Security Law of the People's Republic of China. URL: http://en.npc.gov.cn.cdurl.cn/2021-06/10/c_689311.htm.

⁴ Personal Information Protection Law of the People's Republic of China (Adopted at the 30th Meeting of the Standing Committee of the Thirteenth National People's Congress on August 20, 2021). URL: http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm.

System)¹. Эта система включает три типа платформ: национальную платформу больших данных для государственных вопросов (главный центр управления данными для правительственных операций в Китае), платформу на уровне провинций (охватывает 31 провинцию, автономный район и муниципалитет) и платформу данных для государственных вопросов соответствующих департаментов Госсовета.

Публичные озера данных (цифровые озера) в России

1. Цифровая аналитическая платформа

В 2019 году было отмечено, что многие используемые государственными органами инструменты для статистических исследований несовершенны, в связи с чем в методиках, форматах и показателях наблюдаются разночтения, которые ведут к разнице в интерпретациях. Эту проблему можно решить, переведя статистический учет в цифровой формат².

Национальная программа «Цифровая экономика Российской Федерации» включает в себя федеральный проект «Цифровое государственное управление», предусматривающий создание национальной системы управления данными, одним из компонентов которой должна стать платформа данных. Предполагается, что платформа будет содержать такие элементы, как: национальная система управления данными; единый государственный реестр записей актов гражданского состояния; единый федеральный информационный ресурс, вбирающий сведения о населении Российской Федерации; федеральная государственная информационная система «Единая система нормативной справочной информации».

Создание платформы будет осуществляться на базе Федеральной службы государственной статистики, которой предстоит выполнять функции ее оператора. В 2022 году вступило в силу соответствующее Постановление о государственной информационной системе «Цифровая аналитическая платформа предоставления статистических данных»³.

Система обеспечивает выполнение следующих функций:

– централизованный сбор, хранение, обработка и предоставление информации системы;

– формирование и ведение единого реестра объектов статистического наблюдения;

– формирование и ведение единого реестра форм статистического наблюдения и статистических показателей;

– формирование, ведение и мониторинг выполнения Федерального плана статистических работ;

– защита информации системы, обеспечение целостности и доступности обрабатываемой в ней информации;

– доступ к сведениям, содержащимся в системе, предоставление таких сведений в электронном виде, а также формирование аналитических отчетов на основе информации системы;

– информационное взаимодействие системы с информационными системами субъектов системы, пользователей системы, а также с иными информационными системами;

– разработка (корректировка), согласование и утверждение методик расчета показателей национальных проектов (программ), государственных программ Российской Федерации и их структурных элементов;

– расчет значений показателей национальных проектов (программ), государственных программ Российской Федерации и их структурных элементов, анализ достижения плановых значений указанных показателей.

Система является общедоступной, за исключением информации, доступ к которой ограничен законодательством Российской Федерации.

В 2024 году был принят Закон о цифровой платформе статистических данных⁴. Документ разработан в целях снижения отчетной нагрузки на субъекты малого предпринимательства, оптимизации форм и видов предоставляемой ими отчетности, а также предотвращения ее дублирования различными ведомствами.

2. Национальная система управления данными (НСУД)

Решением Правительства РФ в 2019 году были утверждены Концепция создания и функционирования национальной системы управления данными и План мероприятий («дорожная карта») по созданию национальной системы управления данными на 2019–2021 годы⁵.

¹ China to Build Integrated Govt Big Data System. URL: https://english.www.gov.cn/policies/latestreleases/202210/28/content_WS635b9af9c6d0a757729e1e00.html.

² Концепция создания цифровой аналитической платформы предоставления статистических данных: утв. распоряжением Правительства Рос. Федерации от 17 дек. 2019 г. № 3074-р. URL: <http://static.government.ru/media/files/4YejV8mvcCSeG-WTg2kXprmthtNbWyfrU.pdf>.

³ О государственной информационной системе «Цифровая аналитическая платформа предоставления статистических данных»: Постановление Правительства Рос. Федерации от 22 июня 2021 г. № 956. URL: <http://government.ru/docs/all/135122/>.

⁴ О внесении изменений в статью 5 Федерального закона «О развитии малого и среднего предпринимательства в Российской Федерации» и Федеральный закон «Об официальном статистическом учете и системе государственной статистики в Российской Федерации»: Федер. закон от 22 июля 2024 г. № 206-ФЗ (ред. от 23 мая 2025 г.) // Собр. законодательства Рос. Федерации. 2024. № 31, ст. 4466.

⁵ Концепция создания и функционирования национальной системы управления данными и План мероприятий («дорожная карта») по созданию национальной системы управления данными на 2019–2021 годы: утв. распоряжением Правительства

Под национальной системой управления данными понимается система, состоящая из взаимосвязанных элементов информационно-технологического, организационного, методологического, кадрового и нормативно-правового характера и обеспечивающая достижение целей и выполнение задач, обозначенных в Концепции. При этом государственными данными являются сведения, содержащиеся в информационных ресурсах органов и организаций государственного сектора, а также в информационных ресурсах, созданных в целях реализации полномочий органов и организаций государственного сектора.

Положения Концепции также распространяются на информацию, не являющуюся государственными данными, при ее предоставлении для включения в информационные ресурсы органов и организаций государственного сектора, а также при ее участии в информационном обмене между органами и организациями государственного сектора с учетом требований законодательства Российской Федерации в части соблюдения прав и законных интересов владельцев данных, не являющихся государственными данными, а также соблюдения принципов невмешательства в их хозяйственную деятельность.

В систему входят следующие взаимосвязанные элементы:

1) совокупность нормативных правовых, организационных, методологических правил и процедур, регулирующих отношения органов и организаций государственного сектора, юридических и физических лиц в сфере управления государственными данными, а также обеспечивающих деятельность участников системы;

2) федеральная государственная информационная система «Единая информационная платформа национальной системы управления данными» и иные информационно-технологические элементы системы;

3) цифровая аналитическая платформа предоставления статистических данных.

Обработка сведений, составляющих государственную тайну, в единой информационной платформе не допускается. Исчерпывающий перечень сведений ограниченного доступа, доступ к которым предоставляется посредством единой информационной платформы, определяется в соответствии с законодательством Российской Федерации.

В 2021 году Постановлением Правительства РФ утверждено Положение о федеральной государственной информационной системе «Единая информационная платформа национальной системы управления

данными» (далее – НСУД)¹. НСУД «входит в инфраструктуру, обеспечивающую информационно-технологическое взаимодействие информационных систем, используемых для предоставления государственных и муниципальных услуг и исполнения государственных и муниципальных функций в электронной форме, и создается в целях повышения эффективности обмена и использования государственных данных для предоставления государственных и муниципальных услуг и исполнения государственных и муниципальных функций в электронной форме и иных данных» (п. 3).

НСУД предназначена для решения следующих задач:

- систематизация государственных данных;
- формирование и ведение моделей государственных данных;
- формирование и ведение реестра информационных ресурсов;
- автоматизация процессов управления государственными данными;
- обеспечение контроля качества государственных данных;
- обеспечение сбора, обработки, хранения, предоставления, визуализации и анализа государственных данных из информационных систем органов и организаций государственного сектора и иных данных;
- формирование матрицы доступа.

Участниками НСУД являются органы и организации государственного сектора, а также иные организации, чьи информационные системы участвуют в информационном обмене посредством системы взаимодействия. Органы и организации государственного сектора выступают в роли обладателей государственных данных и пользователей государственных данных. Иные организации, использующие систему взаимодействия, выступают в качестве пользователей государственных данных.

Важно отметить, что уже сейчас НСУД предоставляет по запросу наборы данных для искусственного интеллекта (распознавание такси на изображениях, автотранспортных средств на основе данных БПЛА-съемки; детекция износа и повреждений поверхности дорожного покрытия; распознавание различных транспортных средств для целей определения выбросов загрязняющих веществ в атмосферный воздух от передвижных источников для проведения сводных расчетов загрязнения атмосферного воздуха; распознавание общественного транспорта на изображениях (автобусы, трамваи, троллейбусы) и многое другое).

Рос. Федерации от 3 июня 2019 г. № 1189-р (ред. от 14 мая 2021 г.). URL: <http://static.government.ru/media/files/jYh27VlwiZs44qa0lXlZCa3uu7qqLzl.pdf>.

¹ Об утверждении Положения о федеральной государственной информационной системе «Единая информационная платформа национальной системы управления данными»: Постановление Правительства Рос. Федерации от 14 мая 2021 г. № 733 (в ред. от 28 ноября 2024 г.). URL: <http://static.government.ru/media/files/UN7KKWcK3TgpjAmRcaTzRWUZezblQQWf.pdf>.

3. Цифровое озеро обезличенных персональных данных

В 2024 году Федеральным законом от 8 августа 2024 г. № 233-ФЗ были внесены изменения в Федеральный закон «О персональных данных» и Федеральный закон «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона “О персональных данных”»¹. Закон направлен на обеспечение благоприятных правовых условий для сбора, хранения и обработки данных с использованием новых технологий и вступил в силу 1 сентября 2025 г. Принятие закона призвано существенно повысить эффективность системы защиты прав субъектов персональных данных, а также предоставить им право на использование методологической базы по обезличиванию персональных данных и их передачу в специальную государственную информационную систему (ГИС) для последующего использования.

Закон о персональных данных² дополнен новыми положениями, предусматривающими предоставление операторами персональных данных (ПДн) по требованию Минцифры обезличенных персональных данных в государственную информационную систему министерства, которую определит Правительство РФ. В соответствующем требовании будет указываться перечень данных, которые необходимо передать, а также сроки их предоставления. Минцифры будет формировать из обезличенных данных, сгруппированных по определенному признаку, наборы персональных данных.

При этом установлено, что:

– формирование наборов данных будет осуществляться Минцифры исключительно в случаях, которые определит Правительство РФ;

– формирование наборов данных из специальных категорий персональных данных (есть исключение) и биометрических персональных данных не допускается;

– доступ к дата-сетам будет происходить исключительно в рамках закрытого контура государственной

информационной системы, «вынести» их из нее будет невозможно;

– доступ к сформированным дата-сетам пользователей ГИС – государственных и муниципальных органов и подведомственных им организаций, органов государственных внебюджетных фондов; граждан РФ и российских юридических лиц, сведения о которых внесены в реестр операторов персональных данных и которые соответствуют ряду требований, – будет иметь определенные особенности;

– предоставление результатов обработки составов данных иностранным юридическим лицам, иностранным организациям, не являющимся юридическими лицами, иностранным гражданам и лицам без гражданства по общему правилу запрещается.

Департамент информационных технологий города Москвы будет наделен аналогичными правами по созданию регионального (локального) хранилища обезличенных данных, полученных из государственных и муниципальных информационных систем. Минцифры, в свою очередь, будет иметь право запрашивать обезличенные данные из регионального хранилища (озера данных)³. Закон предусматривает, и это важно отметить, что набор данных, полученных в результате обезличивания, формируется при условии, что последующая обработка этих данных не позволит установить их принадлежность конкретному физическому лицу⁴.

Использование синтетических данных в целях государственного управления

1. Неприкосновенность частной жизни / риски нарушения конфиденциальности

На сегодняшний день этот вопрос является чрезвычайно важным как с точки зрения правового регулирования, так и с точки зрения практической, в части выбора соответствующих технологий и алгоритмов, которые бы обеспечили необходимый уровень анонимности, то есть деперсонализации данных. Современное законодательство о персональных данных называет таковыми любую информацию, которая может быть отнесена к «определенному или определяемому на основании такой информации индивиду»⁵.

¹ О внесении изменений в Федеральный закон «О персональных данных» и Федеральный закон «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона “О персональных данных”»: Федер. закон от 8 авг. 2024 г. № 233-ФЗ // Собр. законодательства Рос. Федерации. 2024. № 33 (ч. 1), ст. 4929.

² О персональных данных: Федер. закон от 27 июля 2006 г. № 152-ФЗ (в ред. от 28 февраля 2025 г.) // Собр. законодательства Рос. Федерации. 2006. № 31 (ч. 1), ст. 3451.

³ С 1 сентября 2025 года компании начнут передавать обезличенные сведения в ГИС. URL: <https://www.garant.ru/news/1749290/>.

⁴ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.172.01.0056.01.ENG.

⁵ Например: О персональных данных: Федер. закон от 27 июля 2006 г. № 152-ФЗ; General Data Protection Regulation. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=legissum:310401_2.

Отсюда возникает вполне закономерный вопрос об алгоритмах и технологиях преобразования реальных данных при создании из них синтетических. Сама технология производства синтетических данных подразумевает как относительно «легкое» преобразование реальных данных (а следовательно, такая информация может содержать отдельные элементы синтетических данных или в дальнейшем приводить к ситуации, когда риск соотнесения ее с тем или иным индивидом объективно сохраняется), так и «полную» их трансформацию (когда полностью синтетические данные повторяют лишь общие шаблоны, рисунки реальных данных) [5].

Эта проблематика относительно давно известна исследователям в области приватности в европейских странах, где также сформировались два процесса преобразования персональных данных для снижения риска их последующего использования: псевдонимизация и анонимизация¹. Первая предполагает в целом сохранение реальных данных, заменяя часть из них псевдонимами и (или) идентификаторами, и допускает обратное преобразование; вторая трансформирует данные более глубоко, в результате чего в идеале утрачивается связь между набором данных и индивидом. Отчасти это напоминает создание синтетических данных. «Легкие» синтетические данные по своему правовому режиму вполне могут рассматриваться как данные, подвергнутые псевдонимизации, и требовать соблюдения общих правил для защиты персональных данных, пусть и на относительно низком уровне с учетом потенциальных рисков для прав индивида. Полностью синтетические данные, с другой стороны, вполне можно рассматривать как анонимизированные данные, то есть как способ полной анонимизации данных [9].

Такая проблематика активно обсуждается и в России, где, как и в других странах, существует реальная потребность бизнеса в получении доступа к большим массивам данных и в выработке условий, при которых данные утрачивают признаки персональных, то есть становятся анонимными. Однако российское законодательство в этой сфере противоречиво, поскольку использует только один вариант описания таких действий – «обезличивание», которое в нормативных документах представляет собой аналог псевдонимизации, или относительно легкую модификацию данных, сохраняющую возможность их обратного преобразования. Этой же позиции придерживается российский регулятор Роском-

надзор, представители которого считают, что обезличенные персональные данные являются по-прежнему персональными данными с сопутствующими рисками необходимости их защиты [9; 2]. Российская практика и исследователи за последнее время неоднократно обосновывали необходимость разработки алгоритмов (правил) гарантированного обезличивания, которое, по сути, является аналогом анонимизации, то есть такого преобразования данных, которое исключает возможность идентификации индивида или установления связи между ним и данными [3].

Аналогичная ситуация и с иной конфиденциальной информацией в части сохранения рисков раскрытия чувствительной информации при использовании «легких» синтетических данных.

2. Риски принятия решений

Использование синтетических данных несет в себе определенные риски в принятии решений на основе указанных данных. Учеными давно отмечается, что сфера государственного управления относительно консервативна к инновациям и требует тщательного выбора средств для своей реализации. Это в особенности касается выбора технологий искусственного интеллекта. В данном случае стоит признать справедливым мнение сторонников установления четкой и последовательной политики в части ограничений на использование и применение технологий генеративного искусственного интеллекта в области публичного управления [7].

С учетом этого вполне очевидной становится проблема проведения испытаний (тестирования), а также оценки применимости технологий искусственного интеллекта для целей публичного управления, вплоть до проведения оценки результатов и алгоритмов искусственного интеллекта экспертами, обладающими необходимыми навыками и компетенциями [18].

Именно в таком случае синтетические данные имеют массу преимуществ, поскольку несут меньшие риски для раскрытия чувствительной информации, но при этом позволяют получить результаты работы алгоритмов генеративного искусственного интеллекта и оценить перспективы и возможности его использования для целей публичного управления. Таким образом, можно привлечь к разработке или оценке искусственного интеллекта частные компании без риска раскрытия конфиденциальной информации или нанесения иного вреда деятельности по публичному управлению.

¹ Pseudonymization according to the GDPR [definitions and examples] // Data Privacy Manager. 2021. Nov. 2. URL: <https://dataprivacymanager.net/pseudonymization-according-to-the-gdpr/>.

Однако следует понимать, что и при использовании синтетических данных присутствуют риски ошибочного принятия решений. Основной проблемой является дилемма между степенью преобразования данных и, как следствие, их точностью, относимостью и рисками раскрытия чувствительной информации. Слишком глубокое и полное преобразование данных при получении полностью синтетических данных несет риски искажения данных, в том числе паттернов данных. В итоге это может отразиться и на процессах принятия решений, которые будут реализовываться с использованием технологий искусственного интеллекта, обученных или построенных на основе таких искаженных искусственных данных. Единственным решением видится достижение определенного баланса между степенью раскрытия чувствительной информации и целями разработки качественных и надежных решений искусственного интеллекта для целей публичного управления, а также проведение тщательной оценки применяемых генеративным искусственным интеллектом синтетических данных [11].

3. Риски технологического суверенитета / риски суверенитета данных

При выстраивании правового регулирования применения синтетических данных важным аспектом могут стать и вопросы обеспечения технологического суверенитета и суверенитета данных. Основная идея создания синтетических данных – обеспечение современной экономики данных важным и необходимым ресурсом. С этой стороны синтетические данные являются ответом на проблематику доступа к чувствительным данным, позволяя получить данные со сходными характеристиками без риска нанесения ущерба правам личности или государства.

Однако современное политическое противостояние государств на международной арене все чаще ставит вопросы обеспечения безопасности и сохранения технологического лидерства, равно как и необходимости основываться на собственных разработках. С учетом этого формируемые государством озера данных и в потенциале полученные на их основе синтетические данные становятся стратегическим ресурсом национальной экономики, доступ к которому требует тщательного и взвешенного подхода.

В сложившейся ситуации стоит обратить внимание на формирующийся подход в отношении предоставления доступа к цифровому озеру обезличенных персональных данных, о котором говорилось в одноименном разделе выше. Согласно правилам, доступ к нему предполагается только для аккредитованных компаний, то есть для российских

компаний, которым государство может доверить работу с такими данными, ожидая определенных позитивных результатов для экономики. Видится вполне объяснимым расширение указанного подхода при обеспечении доступа или использовании национальных озер данных для создания на их основе синтетических данных, а также их последующего реиспользования. Более того, современные российские исследования демонстрируют, что определенную опасность может нести использование данных о характеристиках социальных групп, для которых был предложен термин «данные о человеке». По сути, эта тенденция имеет рациональное обоснование, поскольку данные о характеристиках социальных групп могут быть потенциально использованы для осуществления недружественных действий в отношении государства и (или) указанных социальных групп, что предполагает создание соответствующего охранительного режима.

Синтетические данные с учетом их потенциальных свойств вполне могут характеризовать паттерны социальных групп и категорий индивидов, а также ситуацию в тех или иных отраслях экономики, что, безусловно, ставит вопрос если не об ограничительном режиме, то о необходимости управления и контроля за доступом к таким данным во избежание неблагоприятных последствий для государства [4].

Заключение.

Перспективные направления развития

Цифровое озеро может быть источником информации для обучения генеративного искусственного интеллекта. При этом цифровые озера могут создаваться как коммерческими организациями, так и государственными структурами, которые могут наполнять эти базы данных определенной информацией. Правовой режим использования искусственным интеллектом информации из цифрового озера может быть установлен владельцем цифрового озера либо соответствующим нормативным правовым актом (национальное законодательство), в том числе путем применения экспериментальных правовых режимов в сфере цифровых инноваций.

В этом случае важнейшими условиями получения доступа к такой информации будут являться:

- обеспечение защиты прав и свобод человека и гражданина и защита их персональных данных;
- обеспечение конфиденциальности полученных сведений;
- создание особого доступа к информации иностранных субъектов;
- наличие возможности быстрого прекращения доступа к информации.

Цифровое озеро может заполняться синтетическими данными, когда информация в цифровое озеро поступает от генеративного искусственного интеллекта, который создает новую или преобразует имеющуюся информацию. Важно предусмотреть способы и технические средства, позволяющие идентифицировать реальную и синтетическую информацию, содержащуюся в цифровом озере.

Цифровое озеро может содержать информацию, имеющую побочный характер, – например, изображение автобуса (исходная информация), в котором находятся люди, или человека, переходящего дорогу перед автобусом (побочная информация). Таким образом, искусственный интеллект получает доступ не только к разрешенной информации, но и к сопутствующей информации, что является не совсем корректным.

Наконец, цифровое озеро может быть создано полностью из синтетических данных.

На основании изложенного можно также сделать определенные выводы о перспективах использования формируемых в России озер данных с целью создания синтетических данных для нужд публичного управления и правового регулирования в этой сфере:

1. Использование синтетических данных открывает перспективы для поиска, разработки и оценки применимости технологий искусственного интеллекта в сфере публичного управления. С учетом общей тенденции на повышение «вооруженности» государства технологиями генеративного искусственного интеллекта, обеспечивающими его эффективность, синтетические данные на основе государственных данных способны стать относительно безопасной базой для разработки технологических решений и научных исследований в этой сфере.

2. Правовой режим синтетических данных требует дальнейшего исследования и разработки правовых конструкций. Необходимо как минимум:

– разграничение/соотнесение режимов псевдонимизации, анонимизации и обезличивания (гарантированного обезличивания) данных с алгоритмами и инструментами создания синтетических данных;

– определение параметров охранительного режима синтетических данных, когда их использование может нести риски нанесения ущерба личности или государству;

– разграничение доступа к национальным озерам данных для создания синтетических данных и правил последующего реиспользования синтетических данных.

3. Важным аспектом внедрения синтетических данных в систему публичного управления должна стать процедура и правила оценки и применимости алгоритмов создания синтетических данных для нужд публичного управления в части их точности и соответствия заявленным целям и исключения возможных ошибок и противоречий.

Библиографический список

1. *Филющенко Л. И.* Цифровой профиль работника: проблемы защиты персональных данных // *Законы России: опыт, анализ, практика.* 2024. № 11. С. 93–97.
2. *Куныкина Е.* Обезличенные данные приравнивают к персональным // *Ведомости.* 2021. 12 марта. URL: <https://www.vedomosti.ru/technology/articles/2021/03/11/861164-obezlichennye-dannie>.
3. *Сазонова М.* Обезличивание данных: сохранение баланса между правами граждан и развитием инноваций // *Гарант.ру.* 2021. 24 мая. URL: <https://www.garant.ru/news/1464529/>.
4. *Соловкин С. В.* Публично-правовое регулирование процессов автоматизированного сбора данных о человеке в контексте цифровой трансформации: дис. ... канд. юрид. наук. М., 2025. 278 с.
5. *Abramov M.* Synthetic Data Definition: Pros and Cons. URL: <https://keymakr.com/blog/synthetic-data-definition-pros-and-cons/>.
6. *Bullman H.* The DoD and U.S. Census Bureau Turn to Data Lakes for Modernization and Greater Insights // *Government Technology Insider.* 2020. Aug. 6. URL: <https://governmenttechnologyinsider.com/the-dod-and-u-s-census-bureau-turn-to-data-lakes-for-modernization-and-greater-insights/>.
7. *Bundin M., Martynov A., Aliev Y., Kutuev E.* Legal Aspects of the Use of AI of Public Sector // *Digital Transformation and Global Society (Series: Communications in Computer and Information Science).* Vol. 858. Springer Cham, 2018. Pp. 171–180. DOI: 10.1007/978-3-030-02843-5_14.
8. *Dixon J.* Pentaho, Hadoop, and Data Lakes. URL: <https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes>.
9. *Farrell E. et al.* European Data Spaces – Scientific Insights into Data Sharing and Utilisation at Scale. Luxembourg: Publications Office of the European Union, 2023. 78 p. DOI: 10.2760/400188.
10. *Inmon W. H.* Data Lake Architecture: Designing the Data Lake and Avoiding the Garbage Dump. N.J.: Technics Publications, 2016. 166 p.
11. *Jälmy M., Elvander F., Waterschoot T.* Multi-Channel Low-Rank Convolution of Jointly Compressed Room Impulse Responses // *IEEE Open Journal of Signal Processing.* 2024. Vol. 5. Pp. 850–857. DOI: 10.1109/OJSP.2024.3410089.
12. *Kitchin R.* The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. London: SAGE Publications Ltd., 2014. 240 p. DOI: 10.4135/9781473909472.
13. *Mergel I., Edelmann N., Haug N.* Defining Digital Transformation: Results from Expert Interviews // *Government Information Quarterly.* 2019. Vol. 36. Issue 4. Article 101385. DOI: 10.1016/j.giq.2019.06.002.
14. *Milmo D.* Elon Musk says all human data for AI training “exhausted” // *The Guardian.* 2025. Jan 9. URL: <https://www.theguardian.com/technology/2025/jan/09/elon-musk-data-ai-training-artificial-intelligence>.

15. Pokotylo P. Ethical and Legal Considerations of Synthetic Data Usage. URL: <https://keymakr.com/blog/ethical-and-legal-considerations-of-synthetic-data-usage/>.

16. Stone A. Synthetic Data Supports State and Local Government AI Initiatives // StateTech. 2024. Jul 18. URL: <https://statetechmagazine.com/article/2024/07/synthetic-data-supports-ai-initiatives-for-municipalities-perfcon>.

17. Villalobos P. et al. Will we run out of data? Limits of LLM scaling based on human-generated data // ArXiv. Available at: <https://arxiv.org/abs/2211.04325>. DOI: 10.48550/arXiv.2211.04325

18. Weerts S. Generative AI in public administration in light of the regulatory awakening in the US and EU // Cambridge Forum on AI: Law and Governance. 2025. Vol. 1. Pp. 1–19. DOI: 10.1017/cfl.2024.10.

References

1. Filyushchenko L. I. *Tsifrovoy profil' rabotnika: problemy zashchity personal'nykh dannykh* [Digital Employee Profiling: Problems of Personal Data Protection]. *Zakony Rossii: opyt, analiz, praktika* – Laws of Russia: Experience, Analysis, Practice. 2024. Issue 11. Pp. 93–97. (In Russ.).

2. Kinyakina E. *Obezlichennye dannye priravnayut k personal'nyim* [Depersonalized Data to Be Equated to Personal Data]. *Vedomosti*. 2021, March 12. Available at: <https://www.vedomosti.ru/technology/articles/2021/03/11/861164-obezhichennye-dannie>. (In Russ.).

3. Sazonova M. *Obezlichivanie dannykh: sokhranenie balansa mezhdru pravami grazhdan i razvitiem innovatsiy* [Depersonalization of Data: Maintaining a Balance between the Rights of Citizens and the Development of Innovation]. *Garant*. 2021, May 24. Available at: <https://www.garant.ru/news/1464529/>. (In Russ.).

4. Solovkin S. *Publichno-pravovoe regulirovanie protsessov avtomatizirovannogo sbora dannykh o cheloveke v kontekste tsifrovoy transformatsii: diss. ... kand. yurid. nauk* [Public Law Regulation of Automated Collection of Data on a Person in the Context of Digital Transformation: Cand. jurid. sci. diss.]. Moscow, 2024. 278 p. (In Russ.).

5. Abramov M. Synthetic Data Definition: Pros and Cons. *KeyMakr*. 2024. Available at: <https://keymakr.com/blog/synthetic-data-definition-pros-and-cons/>. (In Eng.).

6. Bullman H. The DoD and U.S. Census Bureau Turn to Data Lakes for Modernization and Greater Insights. *The Government Technology Insider*. 2020. Available at: <https://governmenttechnologyinsider.com/the-dod-and-u-s-census-bureau-turn-to-data-lakes-for-modernization-and-greater-insights/>. (In Eng.).

7. Bundin M., Martynov A., Aliev Y., Kutuev E. Legal Aspects of the Use of AI in Public Sector. *Digital Transformation and Global Society*. Springer, Cham, 2018. Vol. 858. Pp.171–180. (Series: Communications in Computer and Information Science). DOI: 10.1007/978-3-030-02843-5_14. (In Eng.)

8. Dixon J. Pentaho, Hadoop, and Data Lakes. *WordPress.com*. 2010. Available at: <https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes>. (In Eng.).

9. Farrell E., Minghini M. et al. *European Data Spaces – Scientific Insights into Data Sharing and Utilization at Scale*. Luxembourg: Publications Office of the European Union, 2023. 78 p. DOI: 10.2760/400188. (In Eng.).

10. Inmon W. H. *Data Lake Architecture: Designing the Data Lake and Avoiding the Garbage Dump*. New Jersey: Technics Publications, 2016. 166 p. (In Eng.).

11. Jälmy M., Elvander F., Waterschoot T. Multi-Channel Low-Rank Convolution of Jointly Compressed Room Impulse Responses. *IEEE Open Journal of Signal Processing*. 2024. Vol. 5. Pp. 850–857. DOI: 10.1109/OJSP.2024.3410089. (In Eng.).

12. Kitchin R. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: SAGE Publications Ltd., 2014. 240 p. DOI: 10.4135/9781473909472. (In Eng.).

13. Mergel I., Edelman N., Haug N. Defining Digital Transformation: Results from Expert Interviews. *Government Information Quarterly*. 2019. Vol. 36. Issue 4. Article 101385. DOI: 10.1016/j.giq.2019.06.002. (In Eng.).

14. Milmo D. Elon Musk Says All Human Data for AI Training 'Exhausted'. *The Guardian*. 2025. Available at: <https://www.theguardian.com/technology/2025/jan/09/elon-musk-data-ai-training-artificial-intelligence>. (In Eng.).

15. Pokotylo P. Ethical and Legal Considerations of Synthetic Data Usage. *KeyMakr*. 2024. Available at: <https://keymakr.com/blog/ethical-and-legal-considerations-of-synthetic-data-usage/>. (In Eng.).

16. Stone A. Synthetic Data Supports State and Local Government AI Initiatives. *The Statetechmagazine*. 2024. Available at: <https://statetechmagazine.com/article/2024/07/synthetic-data-supports-ai-initiatives-for-municipalities-perfcon>. (In Eng.).

17. Villalobos P., Ho A., Sevilla J., Besiroglu T., Heim L., Hobbhahn M. Will We Run out of Data? Limits of LLM Scaling Based on Human-Generated Data. 2022. *ArXiv*. Available at: <https://arxiv.org/abs/2211.04325>. DOI: 10.48550/arXiv.2211.04325. (In Eng.).

18. Weerts S. Generative AI in Public Administration in Light of the Regulatory Awakening in the US and EU. *Cambridge Forum on AI: Law and Governance*. 2025. Vol. 1. Pp. 1–19. DOI: 10.1017/cfl.2024.10. (In Eng.).

Информация об авторах:

А. В. Мартынов

Доктор юридических наук, профессор,
главный научный сотрудник, заведующий кафедрой
административного и финансового права
Национальный исследовательский Нижегородский
государственный университет им. Н. И. Лобачевского
603022, Россия, г. Нижний Новгород, пр-т Гагарина, 23

ORCID: 0000-0002-6322-818X

ResearcherID: U-7632-2017

Статьи в БД Scopus / Web of Science:

DOI: 10.1007/978-3-319-62530-0_11
DOI: 10.1145/3463677.3463758
DOI: 10.1007/978-3-030-39296-3_21
DOI: 10.1145/3428502.3428531
DOI: 10.1145/3396956.3397866
DOI: 10.1007/978-3-030-13283-5_15
DOI: 10.1109/ICEDEG.2019.8734367
DOI: 10.1007/978-3-030-02843-5_14
DOI: 10.1145/3047273.3047337
DOI: 10.1007/978-3-319-69784-0_20
DOI: 10.1007/978-3-319-49700-6_17
DOI: 10.1145/2912160.2912211
DOI: 10.1145/2910019.2910068
DOI: 10.1145/2757401.2757454
DOI: 10.1145/2691195.2691222
DOI: 10.1007/978-3-030-67238-6_7
DOI: 10.1145/3209415.3209448
DOI: 10.1007/978-3-319-69784-0_12
DOI: 10.1007/978-3-031-04238-6_7
DOI: 10.1109/ICEDEG58167.2023.10121979

Е. В. Ширеева

Кандидат юридических наук, доцент,
старший научный сотрудник, доцент кафедры
административного и финансового права
Национальный исследовательский Нижегородский
государственный университет им. Н. И. Лобачевского
603022, Россия, г. Нижний Новгород, пр-т Гагарина, 23

ORCID: 0000-0002-0373-438X

ResearcherID: U-8242-2017

Статьи в БД Scopus / Web of Science:

DOI: 10.1007/978-3-030-67238-6_7
DOI: 10.1145/3209415.3209448
DOI: 10.1007/978-3-319-69784-0_12
DOI: 10.1007/978-3-031-04238-6_7
DOI: 10.1109/ICEDEG58167.2023.10121979

About the authors:

A. V. Martynov

Lobachevsky State University of Nizhny Novgorod
23, prospekt Gagarina, Nizhny Novgorod, 603022, Russia

ORCID: 0000-0002-6322-818X

ResearcherID: U-7632-2017

Articles in Scopus / Web of Science:

DOI: 10.1007/978-3-319-62530-0_11
DOI: 10.1145/3463677.3463758
DOI: 10.1007/978-3-030-39296-3_21
DOI: 10.1145/3428502.3428531
DOI: 10.1145/3396956.3397866
DOI: 10.1007/978-3-030-13283-5_15
DOI: 10.1109/ICEDEG.2019.8734367
DOI: 10.1007/978-3-030-02843-5_14
DOI: 10.1145/3047273.3047337
DOI: 10.1007/978-3-319-69784-0_20
DOI: 10.1007/978-3-319-49700-6_17
DOI: 10.1145/2912160.2912211
DOI: 10.1145/2910019.2910068
DOI: 10.1145/2757401.2757454
DOI: 10.1145/2691195.2691222
DOI: 10.1007/978-3-030-67238-6_7
DOI: 10.1145/3209415.3209448
DOI: 10.1007/978-3-319-69784-0_12
DOI: 10.1007/978-3-031-04238-6_7
DOI: 10.1109/ICEDEG58167.2023.10121979

E. V. Shireeva

Lobachevsky State University of Nizhny Novgorod
23, prospekt Gagarina, Nizhny Novgorod, 603022, Russia

ORCID: 0000-0002-0373-438X

ResearcherID: U-8242-2017

Articles in Scopus/Web of Science:

DOI: 10.1007/978-3-030-67238-6_7
DOI: 10.1145/3209415.3209448
DOI: 10.1007/978-3-319-69784-0_12
DOI: 10.1007/978-3-031-04238-6_7
DOI: 10.1109/ICEDEG58167.2023.10121979