Научная статья
УДК 343.9
EDN ALFZVS
DOI 10.17150/2500-4255.2026.20(3).242-251

КРИМИНОЛОГИЧЕСКИЕ АСПЕКТЫ ПРИМЕНЕНИЯ СИНТЕТИЧЕСКИХ ДАННЫХ В ГОСУДАРСТВЕННОМ УПРАВЛЕНИИ

А.В. Мартынов, Е.В. Ширеева
Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, г. Нижний Новгород, Российская Федерация

Информация о статье

Дата поступления
22 апреля 2026 г.

Дата принятия к публикации
23 июня 2026 г.

Дата онлайн-размещения
29 июня 2026 г.

Ключевые слова

Генеративный искусственный интеллект; синтетические данные; государственное управление; криминологические риски; дипфейки; системная дискриминация; конфиденциальность

Финансирование

Исследование выполнено за счет гранта Российского научного фонда, проект № 25-28-00491 (https://rscf.ru/project/25-28-00491)

Аннотация. В условиях мировой тенденции, направленной на смягчение правового регулирования и применения генеративного искусственного интеллекта в государственном секторе, важным научным направлением является проведение исследования, посвященного выявлению и установлению криминологических рисков применения генеративного искусственного интеллекта в государственном секторе. Среди таковых в статье отмечаются: создание дипфейков и дезинформирующего контента; предвзятость алгоритмов генеративного искусственного интеллекта и системная дискриминация; проблемы конфиденциальности информации и нарушения права интеллектуальной собственности; совершение преступлений с использованием цифровой личности. В качестве нового цифрового инструмента, направленного на профилактику совершения преступлений с применением генеративного искусственного интеллекта, предлагаются синтетические данные. В статье доказывается возможность нивелирования ими некоторых обозначенных криминологических рисков, поскольку синтетические данные обладают такими преимуществами, как: обеспечивают создание больших объемов высококачественных маркированных данных; снижают предвзятость при автоматизированном принятии управленческого решения; снижают риск раскрытия «чувствительной» и конфиденциальной информации; обеспечивают кибербезопасность; усложняют идентификацию лиц и др. Кроме того, в исследовании предлагаются варианты внедрения синтетических данных в деятельность правоохранительных органов. В частности, отмечается потенциал синтетических данных в моделировании и прогнозировании преступности. Предлагается использовать синтетические данные в целях реализации программы защиты свидетелей. А также отмечается возможность применения анонимизированных и деиндентифицированных синтетических данных, основанных на реальных данных и обладающих высокой точностью информации, при профессиональной подготовке сотрудников правоохранительных органов.

Для цитирования. Мартынов А.В. Криминологические аспекты применения синтетических данных в государственном управлении / А.В. Мартынов, Е.В. Ширеева. — DOI 10.17150/2500-4255.2026.20(3).242-251. — EDN ALFZVS // Всероссийский криминологический журнал. — 2026. — Т. 20, № 3. — С. 242–251.

Original article

CRIMINOLOGICAL ASPECTS OF USING SYNTHETIC DATA IN PUBLIC ADMINISTRATION

Aleksei V. Martynov, Ekaterina V. Shireeva
National Research Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, the Russian Federation

Article info

Received
2026 April 22

Accepted
2026 June 23

Abstract. A global trend for easing the legal regulation and the use of generative artificial intelligence (AI) in the public sector makes it important to conduct research aimed at identifying and establishing the criminological risks of using generative AI in the public sector. Such risks include the creation of deepfakes and disinformation content; biases of generative AI algorithms and systemic discrimination; issues of privacy of information and violations of intellectual property rights; crimes committed with the use of digital identity. Synthetic data are

© Мартынов А.В., Ширеева Е.В., 2026

ISSN 2500-1442

Available online
2026 June 29

Keywords

Generative artificial intelligence; synthetic data; public administration; criminological risks; deepfakes; systemic discrimination; privacy

Funding

The research was funded by the Russian Science Foundation, Project No 25-28-00491 (https://rscf.ru/project/25-28-00491)

presented as a digital instrument aimed at preventing crimes committed with the use of generative AI. The authors demonstrate that they could be used to mitigate some of the abovementioned criminological risks, synthetic data have the following advantages: enable the creation of large volumes of high-quality labeled data; reduce biases in making automated managerial decisions; reduce the risk of compromising "sensitive" and confidential information; ensure cybersecurity; make it harder to identify people, etc. The authors also present some options for the introduction of synthetic data in the work of law enforcement bodies. They specifically note the potential of synthetic data for modelling and predicting crime. It is suggested that synthetic data should be used in the implementation of the witness protection program. It is also highlighted that anonymized and de-identified synthetic data, which is based on real data and characterized by high accuracy, could be used in the professional training of law enforcement employees.

For citation. Martynov A.V., Shireeva E.V. Criminological Aspects of Using Synthetic Data in Public Administration. *Vserossiiskii kriminologicheskii zhurnal = Russian Journal of Criminology*, 2026, vol. 20, no. 3, pp. 242–251. (In Russian). EDN: ALFZVS. DOI: 10.17150/2500-4255.2026.20(3).242-251.

Введение

Применение технологий искусственного интеллекта (ИИ) в сфере государственного управления в Российской Федерации впервые было закреплено на нормативно-правовом уровне в 2019 г.[1] В 2022–2025 гг. произошла «революция» в применении технологий искусственного интеллекта, и в государственный сектор впервые начинают внедряться генеративные модели искусственного интеллекта. Данный этап развития отличается демократизацией технологий ИИ; развитием генеративного ИИ; разработкой открытых моделей ИИ, малых языковых моделей; применением генерации с дополненной выборкой; созданием платформ управления ИИ; разработкой агентского ИИ (ИИ-агентов)[2].

Учеными отмечается, что сама природа генеративного искусственного интеллекта, методы и способы генерации информации имеют как положительные аспекты применения в государственном управлении, так и могут повлечь возникновение правовых рисков в их применении [1–5]. На наш взгляд, важным направлением научного исследования в рассматриваемой сфере представляется выявление правовых рисков криминологического характера, так как совершение именно преступных деяний наносит существенный вред общественным отношениям, складывающимся в государственном секторе, что непосредственно отражается на благополучии граждан, уровне их защищенности и безопасности от пагубного воздействия внедряемых современных технологий. А также эффективность реализации новых внедряемых технологий напрямую зависит от уровня общественного доверия к данной технологии.

Кроме того, выбранный вектор научного исследования обусловлен «мировой тенденцией на смягчение и пересмотр регулирования генеративного искусственного интеллекта. По мнению экспертов АНО «Цифровая экономика», «развитию технологий способствует модель, сочетающая саморегулирование, добровольные стандарты и общие принципы ответственного развития генеративного искусственного интеллекта с детализированными требованиями для чувствительных сфер (госуправление, медицина, финансовые услуги) … В настоящее время большинство стран избегает создания всеобъемлющих, строгих актов (США, Южная Корея, Япония, Великобритания и Россия). Во многих юрисдикциях регулирование закрепляет лишь общие принципы ответственного использования технологий … В России ярко выраженной регуляторной динамики нет — это связано с осознанной политикой властей не создавать избыточные барьеры»[3].

[1] О развитии искусственного интеллекта в Российской Федерации (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») : Указ Президента Российской Федерации от 10 окт. 2019 г. № 490 // Собрание законодательства РФ. 2019. № 41. Ст. 5700.

[2] Доклад Банка России «Применение искусственного интеллекта на финансовом рынке: текущий статус и условия дальнейшего развития». М., 2025. С. 4–8. URL: https://www.cbr.ru/Content/Document/File/185193/Consultation_Paper_20112025.pdf (дата обращения: 10.02.2026).

[3] Интеллект уклоняется от правил. Эксперты проанализировали мировую практику регулирования генеративного ИИ. // Коммерсантъ. URL: https://www.kommersant.ru/doc/8364811 (дата обращения: 22.01.2026).

Криминологические риски применения генеративного искусственного интеллекта в государственном управлении

«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.)» [6, с. 527]. Широкие возможности генеративного искусственного интеллекта по созданию новой информации без участия человека могут применяться в преступных целях и привести к более серьезным и масштабным общественно опасным последствиям.

Во-первых, одним из ключевых криминологических рисков является создание дипфейков и дезинформирующего контента. Как отмечают A.K. Sharma и R. Sharma, «системы генеративного искусственного интеллекта могут генерировать гиперреалистичные изображения, фильмы и аудиоматериалы, которые неотличимы от реальных. Эта функциональность может использоваться для создания дезинформации и дипфейков, что приводит к ряду следующих проблем: подрыв доверия, политические манипуляции, личный вред и клевета, финансовое мошенничество» [1, с. 423]. Тождественной позиции придерживаются Mellouli S., Janssen M., Ojo A., они считают, что «одним из потенциальных рисков применения искусственного интеллекта в государственном секторе является дезинформация и создание фейковых новостей» [2, с. 3].

Одним из резонансных случаев мошенничества в отношении крупной британской энергетической компании в 2019 г. стало использование аудиозаписи с имитацией голоса генерального директора компании, созданного посредством генерации искусственным интеллектом. Компания была обманута на 220 тыс. евро с помощью дипфейка[4]. Кроме того, в 2020 г. во время выборов в США дипфейки использовались для создания провокационных видеороликов с участием политиков и их распространения в социальных сетях. В этих видеороликах содержалась ложная информация, унижающая честь и достоинство кандидатов, участвующих в выборах, которая вводила избирателей в заблуждение[5].

В Российской Федерации один из показательных случаев произошел в 2024 г. в Чебоксарах, когда 75-летней женщине позвонил по видеозвонку в одном из мессенджеров бывший коллега и сообщил, что на ее бывшем рабочем месте проводится служебная проверка и ей скоро позвонят «из правоохранительных органов». В данной мошеннической схеме использовались дипфейки с видеоизображением и имитацией голоса. Последствием такого звонка стал ущерб в 1 млн руб.[6]. В 2025–2026 гг. использовались дипфейки для обмана сотрудников компаний и госслужащих, злоумышленники звонили сотрудникам по видео- или телефонной связи, выдавая себя за руководителя организации и используя технологию дипфейка для подмены лица и голоса[7].

Исходя из приведенных случаев, следует отметить, что с помощью технологии дипфейков и создания дезинформирующего контента могут совершаться преступления, которые имеют широкий круг объектов посягательств. В первую очередь это различного рода мошенничество либо хищение. Органами прокуратуры Российской Федерации разъясняется, что «использование дипфейк-технологий в качестве средства или способа совершения мошенничества возможно квалифицировать как хищение путем ввода, удаления, блокирования, модификации компьютерной информации либо иного вмешательства в функционирование средств хранения, обработки или передачи компьютерной информации или информационно телекоммуникационных сетей, то есть по статье 159.6 УК

[4] UK energy firm probes «deepfake» video of boss // BBC News. 2019. URL: https://www.bbc.com/news/technology-49574808 (дата обращения: 01.02.2026).

[5] Pelosi «drunk» video: Faked footage shows House speaker slurring // BBC News. 2019. URL: https://www.bbc.com/news/world-uscanada-48348059 (дата обращения: 01.02.2026).

[6] Жительница Чебоксар отправила почти миллион рублей мошенникам, использующим для обмана искусственный интеллект // Официальный сайт МВД по Чувашской республике. URL: https://21.xn--b1aew.xn--p1ai/news/item/57232975 (дата обращения: 01.02.2026).

[7] См., например: МВД ИНФОРМИРУЕТ! // Официальный сайт Администрации города Мценска. URL: https://www.adm-mtsensk.ru/ads.html?id=2879 (дата обращения: 01.02.2026) ; МО МВД России «Мценский» информирует!!! // Официальный сайт Администрации Протасовского сельского поселения Мценского района Орловской области. URL: http://xn--80aejtrbcchcncdf.xn--p1ai/mo-mvd-rossii-mtcenskiy-informiruet.html (дата обращения: 01.02.2026).

РФ»[8]. Кроме того, дипфейки применяются с целью совершения преступлений, направленных против конституционных прав и свобод человека и гражданина. Помимо этого, с помощью них могут совершаться преступления против свободы, чести и достоинства личности, в частности, клевета (ст. 128.1 УК РФ), а также такие преступления, направленные против общественной безопасности, как:

– публичное распространение заведомо ложной информации об обстоятельствах, представляющих угрозу жизни и безопасности граждан (ст. 207.1 УК РФ);

– публичное распространение заведомо ложной общественно значимой информации, повлекшее тяжкие последствия (ст. 207.2 УК РФ);

– публичное распространение заведомо ложной информации об использовании Вооруженных Сил Российской Федерации, исполнении государственными органами Российской Федерации своих полномочий, оказании добровольческими формированиями, организациями или лицами содействия в выполнении задач, возложенных на Вооруженные Силы Российской Федерации или войска национальной гвардии Российской Федерации (ст. 207.3 УК РФ).

Следовательно, применение дипфейк-технологий приносит вред и ущерб не только отдельным личностям, а также может привести к более серьезным последствиям, «подстрекающим к подрыву государственной власти, … ставящим под угрозу безопасность и интересы страны, наносящим ущерб репутации страны, подстрекающим к расколу государства, подрывающим национальное единство и социальную стабильность, пропагандирующим терроризм, экстремизм, национальную ненависть, этническую дискриминацию, насилие, непристойность и порнографию, ложную и вредную информацию и другой контент, запрещенный законами и правилами» [3, с. 257–258].

Во-вторых, еще одним криминологическим риском является предвзятость алгоритмов генеративного искусственного интеллекта, системная дискриминация. Ученые полагают, что «при реализации такой функции, как разработка направлений государственной политики и принятие управленческих решений, существуют угрозы смещения в обучающих данных, системная дискриминация» [2, с. 3]. «Системы генеративного искусственного интеллекта обладают потенциалом для преобразования различных областей путем предоставления новой информации и аналитических данных. Однако одна из основных проблем, с которой они сталкиваются, — это возможность воспроизведения и усиления существующих предвзятостей в их обучающих данных. Это может привести к дискриминационным решениям, которые наносят вред отдельным лицам или группам, усугубляя социальное неравенство и предрассудки» [1, с. 426].

Peter Lee отмечает ограничения и риски использования реальных данных для обучения моделей машинного обучения и выделяет среди них предвзятость в автоматизированном принятии решений [4, с. 17]. Исходя из чего, генеративный искусственный интеллект, формируя наборы данных, может создавать предвзятый контент по половой, расовой, национальной, религиозной принадлежности. Ли Яо также считает, что в процессе разработки алгоритмов, подбора обучающих данных, создания и оптимизации моделей и оказания услуг генеративный искусственный интеллект может создавать контент, направленный на дискриминацию по «страновому, географическому, половому, возрастному, профессиональному, медицинскому и другим признакам» [3, с. 258].

Например, в 2025 г. генеративный искусственный интеллект использовался для пропаганды «теории замены» и разжигания национальной ненависти в Европе[9]. Сравнительно недавно в Российской Федерации генеративный искусственный интеллект использовался школьниками из Южно-Сахалинска для создания видео об учителе, разрушающего его репутацию и направленного на дискриминацию педагога по социальному и профессиональному признаку[10].

[8] См., например: Уголовная ответственность за новые преступления, совершаемые с использованием информационно-телекоммуникационных технологий // Официальный сайт Прокуратуры Республики Коми. URL: https://epp.genproc.gov.ru/ru/proc_11/activity/legal-education/explain/e5626152/ (дата обращения: 01.02.2026); Прокуратура Промышленного района г. Владикавказа разъясняет // Официальный сайт муниципального образования г. Владикавказа. URL: https://vladikavkaz-osetia.ru/news/actual.php?ELEMENT_ID=173159 (дата обращения: 01.02.2026).

[9] См., например: Французское информационное агентство // URL: https://www.france24.com/en/live-news/20251013-ai-tools-exploited-for-racist-european-city-videos (дата обращения: 05.02.2026).

[10] Травлю школьных учителей с помощью ИИ-видео расследуют на Сахалине // Информационное агентство «DEITA.RU». URL: https://deita.ru/article/584035 (дата обращения: 18.04.2026).

Таким образом, в рассмотренном разрезе криминологического риска посредством технологий генеративного искусственного интеллекта могут совершаться преступления, направленные на: нарушение равенства прав и свобод человека и гражданина (ст. 136 УК РФ); нарушение права на свободу совести и вероисповедания (ст. 148 УК РФ).

В-третьих, также к криминологическим рискам следует относить проблемы конфиденциальности и нарушения права интеллектуальной собственности при генерации искусственным интеллектом новой информации либо создания нового контента. Эксперты считают, что «системы генеративного искусственного интеллекта часто обучаются на больших наборах данных, содержащих персональную и конфиденциальную информацию. Использование таких данных создает серьезные правовые проблемы, особенно связанные с разрешениями и защитой данных» [1, с. 428].

Итальянские ученые провели исследование по этому вопросу в отношении ChatGPT. По их мнению, «обеспечение безопасности данных становится все более сложной задачей из-за генеративных алгоритмов, которые обрабатывают и собирают огромные объемы данных, преобразуемых посредством токенизации. Токенизация разбивает данные на более мелкие, более управляемые токены, изменяя способ анализа, хранения и использования данных. Хотя токенизация играет важную роль в генеративных моделях, она также создает риски для целостности и конфиденциальности данных» [5, с. 10].

Китайскими учеными отмечается, что «генеративный искусственный интеллект может быть связан с угрозами безопасности данных: сервис генеративного искусственного интеллекта формирует модели, связываясь с большим массивом информации, включая релевантные данные, введенные пользователями. … как только пользователь вводит соответствующие данные с помощью сервиса генеративного ИИ, они становятся частью процесса обучения машинного интеллекта, создавая риск безопасности для личной информации пользователя, конфиденциальной информации, коммерческих тайн и других секретных данных …» [3, с. 250].

Ученые из Канады и Нидерландов, используя свой правовой опыт, приходят к выводу, что еще одним потенциальным риском применения искусственного интеллекта в государственном секторе является нарушение конфиденциальности информации [2, с. 3].

Необходимо также обратить внимание на такое негативное последствие применения генеративного искусственного интеллекта, как массовое нарушение авторских прав, права интеллектуальной собственности, деловой этики [4, с. 17–22]. Алгоритмы, данные и платформы генеративного искусственного интеллекта могут быть использованы для создания монополий и осуществления недобросовестной конкуренции [3, с. 258].

В Российской Федерации фиксируется рост проблем, связанных с неконтролируемым использованием сотрудниками компаний и государственных органов публичных нейросетей (ChatGPT, Gemini и др.) для решения рабочих задач. Аналитики ГК «Солар» при проведении анализа трафиков российских компаний выявили, что «за 2025 год объем конфиденциальной корпоративной информации, отправляемой сотрудниками в публичные нейросети, вырос в 30 раз по сравнению с предыдущим годом. … Сотрудники загружали в нейросети презентации и материалы стратегического планирования, аналитические отчеты и таблицы с бизнес-данными, фрагменты исходного кода, внутреннюю переписку и техническую документацию»[11]. Согласно опросу онлайн-сервиса «Битрикс24» и «Русской школы управления», каждый пятый сотрудник в России признался, что делился с нейросетями конфиденциальными данными. В 16 % компаний такие инциденты случались однократно, еще в 7 % — несколько раз[12]. В США глава американского киберведомства CISA загрузил в публичную версию ChatGPT служебные контракты CISA, которые имели пометку «For Official Use Only» (FOUO) — аналог «для служебного пользования»[13]. Соответственно, любой материал, загруженный в публичный ChatGPT, попадает

[11] Из российских компаний утекло через ИИ-сервисы в 30 раз больше данных, чем в прошлом году // Информационное агентство «CNews». URL: https://www.cnews.ru/news/top/2026-02-04_sotrudniki_rossijskih_kompanij?ysclid=mlj9q8r2qq414777026 (дата обращения 04.02.2026).

[12] С чат-ботами делятся секретами. Сотрудники сами передают конфиденциальную информацию нейросетям // Коммерсантъ. URL: https://www.kommersant.ru/doc/8212875 (дата обращения: 04.02.2026).

[13] Politico // URL: https://www.politico.com/news/2026/01/27/cisa-madhu-gottumukkala-chatgpt-00749361?_bhlid=https://www.csoonline.com/article/4124320/cisa-chief-uploaded-sensitive-government-files-to-public-chatgpt.html (дата обращения: 04.02.2026).

на открытые серверы и может быть использован для ответов на запросы других пользователей.

Безусловно, нарушение конфиденциальности информации в своем большинстве происходит из-за неосторожности самих сотрудников организаций и государственных органов, однако применение технологии генеративного искусственного интеллекта в исследуемом аспекте проблемы гипотетически может привести к совершению преступлений, связанных с нарушением авторских и смежных прав (ст. 146 УК РФ); нарушением правового режима государственной тайны (ст. 283, 283.1, 283.2, 284 УК РФ), служебной тайны (ст. 310 УК РФ), коммерческой и налоговой тайны (ст. 183 УК РФ).

*В-четвертых, новой криминологической угрозой является совершение преступлений с использованием цифровой личности***.** А.М. Кузьмин, Д.А. Свичкарь, П.В. Хенкин вводят дефиницию синтетической цифровой личности и предлагают под ней понимать «цифровую запись о некоторой личности (персоне), содержащую стандартные атрибуты личности (имя, телефон, адрес и т.д.), значения которых полностью сфабрикованы или скомпилированы из реальных и вымышленных данных» [7, с. 251]. Исследователи считают, что для создания синтетический цифровой личности используются основные (ФИО, дата рождения, и уникальные номера официальных документов) и дополнительные (почтовый адрес, номер телефона, IP-адрес, ID устройства) элементы, а также такие методы, как фабрикация, манипуляция, компиляция [7, с. 256].

Однако синтетическую цифровую личность не следует путать с цифровым профилем человека (гражданина). По мнению А.В. Минбалеева, «цифровой профиль — это совокупность актуальных, достоверных данных и иных сведений о гражданах и юридических лицах, формируемых в единой системе идентификации и аутентификации или других информационных систем органов государственной власти и местного самоуправления, а также подведомственных им организаций, взаимодействующих с ней посредством единой системы межведомственного электронного взаимодействия, в целях их предоставления с согласия соответствующих граждан или юридических лиц субъектам, запросившим доступ к этим сведениям посредством инфраструктуры цифрового профиля» [8, с. 111–112]. То есть ключевым отличием синтетической цифровой личности и цифрового профиля человека (гражданина) является реальность и достоверность сведений о человеке.

Необходимо отметить, что синтетическая цифровая личность в понимании вышеуказанных исследователей не имеет ничего общего с синтетическими данными, которые являются предметом исследования настоящей статьи, потому что синтетические данные создаются из синтетических образов, генерируемых специальной компьютерной программой и получаемых оценку со стороны человека для дальнейшего их использования в качестве синтетических данных (информации). Они применяются, например, в целях принятия управленческих решений публично-правовыми образованиями либо они могут использоваться для обучения другого искусственного интеллекта, так как имитируют реальный набор данных без раскрытия конфиденциальной информации.

В свою очередь синтетическая цифровая личность используется преступниками для совершения различного рода мошенничества. По мнению А.М. Кузьмина, Д.А. Свичкарь, П.В. Хенкина, «проблема и масштаб роста мошенничества с использованием синтетической цифровой личности наиболее остро стоит в США … В других странах, в том числе Европы и Азии, мошенничество с синтетической цифровой личностью не выделено в отдельную категорию, и отдельный учет потерь и масштабов подобного мошенничества не ведется … В Российской Федерации в силу принятых законов и выработанных процедур в финансовых организациях проверки персональных данных клиентов аналогичные мошеннические схемы с применением синтетической цифровой личности реализовать гораздо сложнее, но в случае сговора с сотрудником финансовой или другой организации такое становится возможным» [7, с. 259–260].

Синтетические данные как новый цифровой инструмент, направленный на профилактику совершения преступлений с использованием генеративного искусственного интеллекта

В 2024–2025 гг. различные крупные IT-компании мирового уровня заявили об исчерпаемости человеческих ресурсов, используемых в целях обучения искусственного интеллекта, к 2030 г.[14] В связи с чем в качестве альтернативы реальным данным, способным обучать

[14] См., например: Milmo D. Elon Musk says all human data for AI training «exhausted» // Gardian. 2025. URL: https://www.theguardian.com/technology/2025/jan/09/elon-musk-data-ai-training-artificial-intelligence (дата обращения: 10.02.2026).

генеративный искусственный интеллект, IT-специалисты предложили применять синтетические данные.

С точки зрения правовой сферы, синтетические данные — это «сгенерированные специальной компьютерной программой синтетические образы данных (информации), получившие оценку со стороны человека для дальнейшего их использования в качестве синтетических данных (информации), которые формируются посредством модуля генерации набора данных, применяющего математические методы для аугментации данных и использующего для получения нового набора данных нейронные сети» [9, с. 143]. Синтетические данные характеризуются такими идентификационными признаками, как:

– «создаются компьютерными программами генеративного искусственного интеллекта;

– основным методом получения синтетических данных является глубокое машинное обучение, основанное на языковых и графических моделях;

– источником получения исходных, промежуточных и итоговых данных являются нейросети;

– синтетические данные по свойствам приближены к реальным данным;

– невозможно обычными способами (в том числе человеку) отличить синтетические данные от реальной информации;

– синтетические данные являются новой информацией (данными), которые ранее не были созданы кем-либо;

– синтетические данные следует считать единственно возможной технологией дальнейшего развития генеративного искусственного интеллекта при существующем ограничении доступа к информации и данным (их исчерпаемости);

– в отличие от реальной информации существует неисчерпаемость синтетических данных;

– синтетические данные (информация) формируются из синтетических образов данных (информации), которые получают оценку со стороны человека для дальнейшего их использования в качестве таковых;

– юридическая ответственность за генерацию синтетических данных может быть возложена на разработчиков компьютерных программ либо пользователей компьютерных программ» [там же, с. 143–144].

Peter Lee отмечает следующие преимущества применения синтетических данных: обеспечение создания больших объемов высококачественных маркированных данных, снижение предвзятости при автоматизированном принятии решений и предотвращение нарушения авторских прав [4, с. 27–34].

P. Calcraft, I. Thomas, M. Maglicic, A. Sutherland полагают, «поскольку записи в наборе синтетических данных являются искусственными и не соответствуют записям в реальных данных, при работе с синтетическими данными, в принципе, существует меньше опасений относительно риска раскрытия информации» [10, с. 7]. Ими приводится пример, что Министерство обороны Великобритании «рассматривает возможность создания реалистичных синтетических данных, пригодных для предоставления внешним исследователям, желающим провести анализ данных в случаях, когда реальные данные могут раскрыть конфиденциальную информацию, например, местоположение или производительность определенных видов оборудования» [там же, с. 12]. Кроме того, специалистами Исследовательского отделения Центра системных лонгитюдных данных Мэриленда доказано, «что полностью синтетическая структура данных отлично справляется с защитой как от индивидуального риска раскрытия, так и от риска раскрытия атрибутов» [11, с. 13].

M.S. Gal, O. Lynskey считают, что «синтетические данные используются для повышения конфиденциальности или кибербезопасности, тем самым позволяя более широко использовать ценные данные для исследований и принятия решений» [12, с. 1090].

По мнению K. Jenkins, «синтетические данные иногда могут быть более точными, чем реальные данные, когда набор данных реального мира содержит известные источники ошибок, а синтетические данные корректируются для их устранения» [13, с. 32]. Исходя из чего, синтетические данные позволяют нивелировать проблему предвзятости алгоритмов генеративного искусственного интеллекта и системной дискриминации. Данная мысль подтверждается также проведенным экспериментом, который выявил, что синтетические данные «позволяют анализировать данные реального мира, выявлять и компенсировать смещение, а также генерировать гораздо более крупные наборы данных, которые лучше подходят для небольших групп» [там же].

Среди ключевых преимуществ синтетических данных S. Kurapati, L. Gilli выделяют следующие: они могут быть эффективным механизмом

защиты от прямой повторной идентификации; они обеспечивают дополнительные гарантии конфиденциальности, такие как метрики оценки риска конфиденциальности; скрывая индивидуальные данные в статистических свойствах данных, синтетические данные обеспечивают более точное представление сложных наборов данных, одновременно защищая идентификацию отдельных лиц; они могут помочь устранить несбалансированность или предвзятость наборов данных (соблюсти принцип недискриминации) [14, с. 8–9].

Таким образом, многообразие приведенных научных мнений позволяет обобщить и выявить такие преимущества применения синтетических данных в государственном секторе, как:

– обеспечение создания больших объемов высококачественных маркированных данных;

– снижение предвзятости при автоматизированном принятии управленческого решения;

– снижение риска раскрытия «чувствительной» и конфиденциальной информации;

– обеспечение кибербезопасности;

– усложнение идентификации лиц и др.

Заключение

Исходя из выделенных преимуществ синтетических данных, особенностей их генерации и характерных черт, можно сформулировать вывод о том, что они позволяют нивелировать, как минимум, два рассмотренных в статье криминологических риска: во-первых, связанных с предвзятостью алгоритмов генеративного искусственного интеллекта, системной дискриминацией; а во-вторых, с разрешением проблемы конфиденциальности информации и нарушения права интеллектуальной собственности.

Однако также необходимо отметить потенциал синтетических данных в моделировании и прогнозировании преступности. Синтетические данные позволяют разрабатывать и тестировать высокоточные программы, направленные на создание различного рода сценариев совершения преступлений с большим количеством вводимых условий и обстоятельств, при этом имея низкий уровень риска незаконного получения доступа к реальным данным. Например, международной исследовательской группой был создан «PLUS — полуавтоматизированный конвейер для обнаружения мошенничества в тендерной документации. Он включает эвристический метаклассификатор для тендерной документации и модуль качества данных. Оба модуля демонстрируют многообещающие результаты после проверки концепции, подтверждая актуальность PLUS для автоматизации расследования процесса торгов. Также разработчики предоставляют два модуля, основанных на реальных данных без раскрытия чувствительной информации: построение журналов аудита для обнаружения мошенничества и базу данных цен для обнаружения завышения цен» [15, с. 1].

Кроме того, представляется возможным использование синтетических данных в Российской Федерации в целях реализации программы защиты свидетелей. В частности, синтетические данные снизят риски утечки информации, содержащей личные сведения о человеке.

И наконец, анонимизированные и деиндентифицированные синтетические данные, основанные на реальных данных, которые обладают высокой точностью информации, можно использовать при профессиональной подготовке сотрудников правоохранительных органов.

В заключение хотелось бы отметить, что в настоящее время в Минцифры рассматриваются предложения, направленные Ассоциацией больших данных (объединяет «Яндекс», VK, «Сбер», Газпромбанк, Т-Банк, Россельхозбанк, банк «Точка», «Мегафон», «Ростелеком», Билайн, МТС и др.), о повышении доступности данных для обучения ИИ при соблюдении прав граждан на неприкосновенность частной жизни; ускорении внедрения технологий повышения конфиденциальности для обработки чувствительных данных; безопасном распределении данных между более широким кругом экономических агентов; увеличении объемов и разнообразия данных, доступных для обучения моделей ИИ; появлении новых видов инновационных продуктов, услуг и сервисов на основе данных. Помимо прочего, Ассоциацией предлагается создать правовые условия для свободного перетока данных между различными отраслями и секторами экономики, сформировать безопасную, открытую и конкурентную систему обмена данными[15]. Это означает, что вопрос внедрения синтетических данных в различные сферы государственного управления является перспективой ближайшего будущего.

[15] Это база: какие поправки в законодательство предложил бизнес для развития ИИ // Forbes. URL: https://www.forbes.ru/tekhnologii/553967-eto-baza-kakie-popravki-v-zakonodatel-stvo-predlozil-biznes-dla-razvitia-ii (дата обращения: 22.01.2026).

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Sharma A.K. Generative Artificial Intelligence and Legal Frameworks: Identifying Challenges and Proposing Regulatory Reforms / A.K. Sharma, R. Sharma. — DOI 10.17803/2713-0533.2024.3.29.415-451. — EDN HRTCGK // Kutafin Law Review. — 2024. — № 11. — P. 415–451.

2. Mellouli S. Introduction to the Issue on Artificial Intelligence in the Public Sector: Risks and Benefits of AI for Governments / S. Mellouli, M. Janssen, A. Ojo. — DOI 10.1145/3636550 // Digital Government: Research and Practice. — 2024. — Vol. 5, № 1. — P. 1–6.

3. Ли Яо. Нормативно-правовое регулирование генеративного искусственного интеллекта в Великобритании, США, Европейском союзе и Китае / Яо Ли. — DOI 10.17323/2072-8166.2023.3.245.267. — EDN YITZOA // Право. Журнал Высшей школы экономики. — 2023. — Т. 16, № 3. — С. 245–267.

4. Lee P. Synthetic Data and the Future of AI / P. Lee // Cornell Law Review (Forthcoming). — 2024. — Vol. 110, no. 1.

5. Cordella A. Regulating Generative AI: The Limits of Technology-Neutral Regulatory Frameworks. Insights from Italy's Intervention on ChatGPT / A. Cordella, F. Gualdi. — DOI 10.1016/j.giq.2024.101982 // Government Information Quarterly. — 2024. — № 41. — P. 101982.

6. Мартынов А.В. Правовые и практические вопросы использования синтетических данных для целей государственного управления в России / А.В. Мартынов, Е.В. Ширеева. — DOI 10.17072/1995-4190-2025-70-526-538. — EDN PNOBJC // Вестник Пермского университета. Юридические науки. — 2025. — № 4. — С. 526–538.

7. Кузьмин А.М. Мошенничество с использованием синтетических цифровых личностей / А.М. Кузьмин, Д.А. Свичкарь, П.В. Хенкин. — DOI 10.25559/sitito.019.202302.251-261. — EDN UOKXUQ // Современные информационные технологии и ИТ-образование. — 2023. — Т. 19, № 2. — С. 251–261.

8. Минбалеев А.В. Понятие и правовая природа цифрового профиля человека / А.В. Минбалеев. — DOI 10.14529/law220117. — EDN MMFGYL // Вестник Южно-Уральского государственного университета. Серия: Право. — 2022. — Т. 22, № 1. — С. 110–116.

9. Мартынов А.В. Понятие и виды синтетических данных: к вопросу о необходимости правового регулирования в российском законодательстве / А.В. Мартынов, Е.В. Ширеева. — DOI 10.61205/S160565900035508-2. — EDN HQWCPL // Журнал российского права. — 2025. — Т. 29, № 11. — С. 135–148.

10. Accelerating Public Policy Research with Synthetic Data / P. Calcraft, I. Thomas, M. Maglicic, A. Sutherland. — ADR UK, 2021. — 42 p.

11. Expanding MLDS Data Access and Research Capacity with Synthetic Data Sets / M.E. Woolley, L.M. Stapleton, R. Goldstein [et al.]. — Baltimore : Maryland Longitudinal Data System Center, 2020. — 20 p.

12. Gal M.S. Synthetic Data: Legal Implications of the Data-Generation Revolution / M.S. Gal, O. Lynskey // Iowa Law Review. — 2023. — Vol. 109. — P. 1087–1156.

13. Jenkins K. Synthetic Data and Public Policy. Supporting Real-World Policymakers with Algorithmically Generated Data / K. Jenkins. — DOI 10.26686/pq.v19i2.8234 // Policy Quarterly. — 2023. — Vol. 19, no. 2. — P. 29–39.

14. Kurapati S. Synthetic Data: A Convergence between Innovation and GDPR / S. Kurapati, L. Gilli. — DOI 10.63567/gk6xj346 // Journal of Open Access to Law. — 2023. — Vol. 11. — P. 1–12.

15. PLUS: A Semi-automated Pipeline for Fraud Detection in Public Bids / M.A. Brandão, A.P.G. Reis, B.M.A. Mendes [et al.]. — DOI 10.1145/3616396 // Digital Government: Research and Practice. — 2024. — Vol. 5, no. 1. — P. 5–16.

REFERENCES

1. Sharma A.K., Sharma R. Generative Artificial Intelligence and Legal Frameworks: Identifying Challenges and Proposing Regulatory Reforms. *Kutafin Law Review*, 2024, no. 11 (3), pp. 415–451. EDN: HRTCGK. DOI: 10.17803/2713-0533.2024.3.29.415-451.

2. Mellouli S., Janssen M., Ojo A. Introduction to the Issue on Artificial Intelligence in the Public Sector: Risks and Benefits of AI for Governments. *Digital Government: Research and Practice*, 2024, vol. 5, no. 1, pp. 1–6. DOI: 10.1145/3636550.

3. Li Yao. Specifics of Regulatory and Legal Regulation of Generative Artificial Intelligence in the UK, USA, EU and China. Pravo. *Zhurnal Vysshey shkoly ekonomiki = Law. Journal of the Higher School of Economics*, 2023, vol. 16, no. 3, pp. 245–267. (In Russian). EDN: YITZOA. DOI: 10.17323/2072-8166.2023.3.245.267.

4. Lee P. Synthetic Data and the Future of AI. *Cornell Law Review (Forthcoming)*, 2024, vol. 110, no. 1.

5. Cordella A., Gualdi F. Regulating Generative AI: The Limits of Technology-Neutral Regulatory Frameworks. Insights from Italy's Intervention on ChatGPT. *Government Information Quarterly*, 2024, no. 41, pp. 101982. DOI: 10.1016/j.giq.2024.101982.

6. Martynov A.V., Shireeva E.V. Legal and Practical Issues of Synthetic Data Use for Public Administration Purposes in Russia. *Vestnik Permskogo universiteta. Yuridicheskie nauki = Perm University Herald. Juridical Science*, 2025, no. 4, pp. 526–538. (In Russian). EDN: PNOBJC. DOI: 10.17072/1995-4190-2025-70-526-538.

7. Kuzmin A.M., Svichkar D.A., Khenkin P.V. Synthetic Identity Fraud. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie = Modern Information Technology and IT-Education*, 2023, vol. 19, no. 2, pp. 251–261. (In Russian). EDN: UOKXUQ. DOI: 10.25559/sitito.019.202302.251-261.

8. Minbaleev A.V. The Concept and Legal Nature of the Digital Profile of a Person. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Pravo = Bulletin of South Ural State University. Series: Law*, 2022, vol. 22, no. 1, pp. 110–116. (In Russian). EDN: MMFGYL. DOI: 10.14529/law220117.

9. Martynov A.V., Shireeva E.V. The Concept and Types of Synthetic Data: On the Issue of the Necessity of Legal Regulation in Russian Legislation. *Zhurnal rossiyskogo prava = Russian Law Journal*, 2025, vol. 29, no. 11, pp. 135–148. (In Russian). EDN: HQWCPL. DOI: 10.61205/S160565900035508-2.

10. Calcraft P., Thomas I., Maglicic M., Sutherland A. *Accelerating Public Policy Research with Synthetic Data*. ADR UK, 2021. 42 p.

11. Woolley M.E., Stapleton L.M., Goldstein R., Bonnéry D., Lachowicz M. *Expanding MLDS Data Access and Research Capacity with Synthetic Data Sets*. Baltimore, Maryland Longitudinal Data System Center, 2020. 20 p.

12. Gal M.S., Lynskey O. Synthetic Data: Legal Implications of the Data-Generation Revolution. *Iowa Law Review*, 2023, vol. 109, pp. 1087–1156.

13. Jenkins K. Synthetic Data and Public Policy. Supporting Real-World Policymakers with Algorithmically Generated Data. *Policy Quarterly*, 2023, vol. 19, no. 2, pp. 29–39. DOI: 10.26686/pq.v19i2.8234.

14. Kurapati S., Gilli L. Synthetic Data: A Convergence Between Innovation and GDPR. *Journal of Open Access to Law*, 2023, vol. 11, pp. 1–12. DOI: 10.63567/gk6xj346.

15. Brandão M.A., Reis A.P.G., Mendes B.M.A., Bacha C.A., Oliveira G.P. PLUS: A Semi-Automated Pipeline for Fraud Detection in Public Bids. *Digital Government: Research and Practice*, 2024, vol. 5, no. 1, pp. 5–16. DOI: 10.1145/3616396.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Мартынов Алексей Владимирович — доктор юридических наук, профессор, главный научный сотрудник, заведующий кафедрой административного и финансового права юридического факультета, Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского, 603022, Российская Федерация, г. Нижний Новгород, пр-кт Гагарина, 23, docpred@yandex.ru.

Ширеева Екатерина Валерьяновна — кандидат юридических наук, доцент, старший научный сотрудник, доцент кафедры административного и финансового права юридического факультета Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского, 603022, Российская Федерация, г. Нижний Новгород, пр-кт Гагарина, 23, shireevaekaterina@yandex.ru.

INFORMATION ABOUT THE AUTHORS

Martynov, Aleksei V. — Doctor of Law, Professor, Chief Researcher, Head, Department of Administrative and Financial Law, Law Faculty, National Research Lobachevsky State University of Nizhny Novgorod, 23 Gagarin Ave., Nizhny Novgorod, 603022, the Russian Federation, docpred@yandex.ru.

Shireeva, Ekaterina V. — Ph.D. in Law, Ass. Professor, Senior Researcher, Department of Administrative and Financial Law, Law Faculty, National Research Lobachevsky State University of Nizhny Novgorod, 23 Gagarin Ave., Nizhny Novgorod, 603022, the Russian Federation, shireevaekaterina@yandex.ru.

ВКЛАД АВТОРОВ

Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

CONTRIBUTION OF THE AUTHORS

The authors contributed equally to this article. The authors declare no conflicts of interests.