ФЕДЕРАЛЬНОЕ АГЕНТСТВО
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПНСТ

СИНТЕЗ ДАННЫХ
Архитектура процесса синтеза данных. Методы синтеза

Издание официальное

Москва
Российский институт стандартизации
2026

Предисловие

1 РАЗРАБОТАН Некоммерческой организацией «Ассоциация больших данных» (НО «АБД») при участии Автономной некоммерческой организации «Национальный технологический центр цифровой криптографии» (АНО «НТЦ ЦК»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от__ _____ 202_ г. № ____

Правила применения настоящего стандарта и проведения его мониторинга установлены в ГОСТ Р 1.16—2011 (разделы 5 и 6).

Федеральное агентство по техническому регулированию и метрологии собирает сведения о практическом применении настоящего стандарта. Данные сведения, а также замечания и предложения по содержанию стандарта можно направить не позднее, чем за 4 мес до истечения срока его действия разработчику настоящего стандарта по адресу: 123112 Москва, Пресненская набережная, д. 10 стр. 2 и/или в Федеральное агентство по техническому регулированию и метрологии по адресу: 123112 Москва, Пресненская набережная, д. 10, стр.2.

В случае отмены настоящего стандарта соответствующая информация будет опубликована в ежемесячном информационном указателе «Национальные стандарты» и также будет размещена на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (www.rst.gov.ru)

Содержание

Введение

Эффективность синтеза данных тесно связана с качеством процессов синтеза и архитектуры используемых систем. Синтетические данные применяются в различных практических сценариях, таких как тестирование информационных систем, обучение моделей искусственного интеллекта (ИИ), обмен данными с контрагентами и обогащение существующих наборов данных. Многие из этих сценариев требуют специфического подхода к качеству и конфиденциальности синтетических данных, что делает критически важным детальное описание процессов и архитектуры системы синтеза.

Для тестирования информационных систем синтетические данные должны точно отражать структуру и особенности реальных данных, обеспечивая при этом предотвращение потенциальной утечки данных (полностью синтетические данные). Обучение моделей требует от синтетических данных репрезентативности и достоверности, что позволяет моделям эффективно адаптироваться к задачам реального мира. При обмене данными с контрагентами особое внимание должно уделяться доверенным технологиям и методам синтеза, чтобы минимизировать вероятность утечки данных. Для задач обогащения данных при интеграции синтетические данные должны сохранять целостность и совместимость с другими наборами данных, обеспечивая при этом их полезность для дальнейшего анализа.

Таким образом, архитектура и процессы синтеза данных должны быть гибкими и адаптивными, чтобы соответствовать разнообразным требованиям различных сценариев применения. Достоверность и полнота синтетических данных, их соответствие требованиям к доверенным технологиям ИИ, полезность получаемых данных

зависят не только от алгоритмов синтеза, но и от порядка их применения, организации хранения данных и управления вероятностями утечки данных на каждом этапе.

Цель настоящего стандарта — предложить эталонное описание процессов и архитектуры системы синтеза данных, что необходимо для достижения высоких показателей качества и конфиденциальности данных в различных контекстах их использования. Стандарт служит ориентиром для разработки, внедрения и эксплуатации систем синтеза данных, соответствующих требованиям конфиденциальности и качества, что особенно важно в условиях возрастающих требований к конфиденциальности и управлению данными.

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

СИНТЕЗ ДАННЫХ

Архитектура процесса синтеза данных.

Методы синтеза

Data synthesis. Data synthesis process architecture. Synthesis methods

Срок действия — с __—____—____
до __—____—____

1 Область применения

Настоящий стандарт устанавливает процесс синтеза данных (полностью синтетические данные) для систем искусственного интеллекта (ИИ) и архитектуру соответствующей системы, включая основные компоненты, процессы генерации данных и интеграцию с системами ИИ.

Настоящий стандарт устанавливает требования к выполнению каждого из этих этапов с учетом специфики обработки данных, мониторинга соответствия и управления процессом синтеза доверенным технологиям синтеза, а также поддержке уровня приватности на всех стадиях жизненного цикла синтетических данных.

Стандарт применим к организациям любых форм собственности и масштабов, использующим технологии ИИ и синтетические данные.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р 34.12 Информационная технология. Криптографическая защита информации. Блочные шифры

Издание официальное

пнст

ГОСТ Р 57098 Системная и программная инженерия. Управление жизненным циклом. Руководство для описания процесса

ГОСТ Р 57100 Системная и программная инженерия. Описание архитектуры

ГОСТ Р 57640/ISO/IEC TS 33052:2016 Информационные технологии. Эталонная модель процесса (ЭМП) для управления информационной безопасностью

ГОСТ Р 57193–2025 Системная и программная инженерия. Процессы жизненного цикла систем

ГОСТ Р 59277 Системы искусственного интеллекта. Классификация систем искусственного интеллекта

ГОСТ Р 59330 Системная инженерия. Защита информации в процессе управления моделью жизненного цикла системы

ГОСТ Р 59347 Системная инженерия. Защита информации в процессе определения архитектуры системы

ГОСТ Р 59898 Оценка качества систем искусственного интеллекта. Общие положения

ГОСТ Р 59926/ISO/IEC TR 20547-2:2018 Информационные технологии. Эталонная архитектура больших данных. Часть 2. Варианты использования и производные требования

ГОСТ Р 59992 Системная инженерия. Системный анализ процесса управления моделью жизненного цикла системы

ГОСТ Р 70466/ISO/IEC TR 20547-1:2022 Информационные технологии. Эталонная архитектура больших данных. Часть 1. Структура и процесс применения

ГОСТ Р 70889 (ИСО/МЭК 8183:2023) Информационные технологии. Искусственный интеллект. Структура жизненного цикла данных

ГОСТ Р 71440 Информационные технологии. Оценка процессов. Руководство по определению рисков в процессах

ГОСТ Р 71539–2024 (ИСО/МЭК 5338:2023) Искусственный интеллект. Процессы жизненного цикла системы искусственного интеллекта

ГОСТ Р ИСО 9000 Системы менеджмента качества. Основные положения и словарь

ГОСТ Р ИСО/МЭК 12207 Информационная технология. Системная и программная инженерия. Процессы жизненного цикла программных средств

ГОСТ Р ИСО/МЭК 20547-3 Информационные технологии. Эталонная архитектура больших данных. Часть 3. Эталонная архитектура

ГОСТ Р ИСО/МЭК 24668 Информационные технологии. Искусственный интеллект. Структура управления процессами аналитики больших данных

ГОСТ Р ИСО/МЭК 33001 Информационные технологии. Оценка процесса. Понятия и терминология

ПНСТ 1064—2026 Синтез данных. Термины и определения. Основные положения

ПНСТ 1066—2026 Синтез данных. Описание результатов процесса синтеза. Методика оценки качества

П р и м е ч а н и е — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования – на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется

использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящем стандарте применены термины по ПНСТ 1064—2026, а также следующие термины:

3.1 **генератор синтетических данных** (data synthesis generator): Реализация метода создания синтетических данных, включая модели машинного обучения, статистические методы или подходы, основанные на правилах.

Примечание — Генератор синтетических данных на основе математических и алгоритмических методов [например, генеративно-состязательных сетей или вариационных автокодировщиков, а также другие современные архитектуры генеративных моделей, включая диффузионные модели (Diffusion Models) и авто-регрессивные трансформеры] генерирует искусственные данные, которые могут использоваться для обучения, тестирования моделей машинного обучения или других целей, связанных с анализом данных, при этом обеспечивая снижение вероятности утечки данных.

3.2 **единица приватности** (unit of privacy): Минимальный логический элемент данных, относительно которого обеспечивается приватность при применении методов защиты, таких как дифференциальная приватность.

Примечание – Единицей приватности может быть, например, отдельная запись в наборе данных, индивидуальный субъект (пользователь), событие или транзакция — в зависимости от контекста применения и модели угроз.

3.3 **статистическая близость** (statistical proximity): Мера, описывающая, насколько распределение синтетических данных близко к распределению исходных реальных данных.

Примечание — Высокая статистическая близость указывает на то, что синтетические данные сохраняют ключевые статистические характеристики реальных данных, что делает их пригодными для аналитики и моделирования.

3.4 **уровень приватности данных** (data privacy level): Количественная или качественная оценка степени защищенности обработанных данных от утечки данных, полученная в результате формализованной процедуры анализа.

3.5 **чувствительные данные** (sensitive data): Сведения, несанкционированное раскрытие или модификация которых может повлечь за собой негативные последствия, включая финансовые, репутационные или правовые риски, для субъектов данных, организаций или общества в целом.

Примечание — Исходные данные могут иметь разный уровень чувствительности, в зависимости от объема и характера сведений, которые они содержат.

4 Сокращения

В настоящем стандарте применены следующие сокращения:

ЖЦ — жизненный цикл;

ИИ — искусственный интеллект;

ИТ — информационные технологии;

МО — машинное обучение;

ПДн — персональные данные;

ФЛ — физическое лицо;

PRM — эталонная модель процесса (process reference model);

GAN — генеративно-состязательная сеть (generative adversarial networks);

RDP — дифференциальная приватность Реньи (rényi differential privacy).

5 Обзор жизненного цикла синтеза

5.1 Общие положения

Синтетические данные могут использоваться для широкого круга задач в качестве решения, когда публикация исходных данных невозможна из-за ограничений законодательства или когда для задач тренировки моделей ИИ необходимы дополнительные данные. Синтетические данные могут создаваться различными способами. В настоящем стандарте рассматриваются методы создания полностью синтетических данных с использованием систем ИИ, которые на основе извлеченных статистических закономерностей и других особенностей исходных данных синтезируют искусственные данные. Следуя классификации ГОСТ Р 59277, системы синтеза могут быть отнесены к автоматизированным системам с обратной связью.

Системы ИИ в процессе создания, эксплуатации и вывода из эксплуатации проходят стадии ЖЦ, характерные для любой информационной системы, определенные ГОСТ Р ИСО/МЭК 12207. ЖЦ систем ИИ, работающих с данными, тесно связан с ЖЦ данных, которые поддерживают ИИ. Эти циклы зависят друг от друга, так как данные проходят этапы сбора, обработки, хранения и использования параллельно с разработкой и эксплуатацией системы ИИ.

Синтез данных является частным случаем процессов, описанных в ГОСТ Р 70889, и включает специфические уточнения и коррекции, характерные для генеративного ИИ. При описании ЖЦ синтетических данных в настоящем стандарте используются следующие ключевые понятия, сформулированные и уточненные на основе определений, приведенных в ПНСТ 1064—2026:

- система синтеза представляет собой совокупность инструментов и алгоритмов, реализованных в виде информационной системы, обеспечивающей подготовку данных, построение статистических моделей, генерацию синтетических данных и контроль уровня доверия к технологиям синтеза;
- модель синтеза (генератор синтетических данных) в данном контексте трактуется как разновидность модели МО, способная извле-

кать закономерности из исходных данных и использовать их для генерации новых, синтетических записей на основе статистических зависимостей (см. ПНСТ 1064—2026, пункт 3.1.1);

- синтетические данные определяются как искусственно созданные данные, воспроизводящие структуру и статистические свойства реальных данных, но не содержащие информацию, позволяющую напрямую соотнести их с реальными объектами, включая ПДн. Такие данные применяются для анализа, тестирования и обучения моделей при сохранении требований к приватности (см. ПНСТ 1064—2026, пункт 3.1.3).

Каждое из этих понятий обладает своим ЖЦ, который должен быть интегрирован в общий процесс синтеза данных. Для определения стадий ЖЦ используется системный анализ, определенный в ГОСТ Р 59992. Полный цикл синтеза данных включает следующие обязательные стадии:

- стадия 1 «Анализ»;
- стадия 2 «Формирование требований»;
- стадия 3 «Планирование работы с синтетическими данными»;
- стадия 4 «Сбор и агрегация эталонных данных»;
- стадия 5 «Подготовка эталонных данных»;
- стадия 6 «Подстройка параметров модели»;
- стадия 7 «Обучение модели»;
- стадия 8 «Синтез выходных данных»;
- стадия 9 «Постпроцессинг синтетических данных»;
- стадия 10 «Валидация и оценка качества синтетических данных»;
- стадия 11 «Публикация синтетических данных»;
- стадия 12 «Развертывание модели искусственного интеллекта»;
- стадия 13 «Эксплуатация системы искусственного интеллекта»;
- стадия 14 «Вывод синтетических данных из эксплуатации»;
- стадия 15 «Вывод из эксплуатации системы».

Рекомендации по организации ЖЦ синтеза приведены в приложении А.

5.2 Стадии жизненного цикла синтеза

5.2.1 Стадия 1 «Анализ»

Стадия 1 «Анализ» включает в себя задачи по выявлению потребности в синтетических данных, определению целей и задач синтеза на основе имеющейся потребности. На данной стадии выясняется, существуют ли необходимые источники данных, кто является потребителем синтетических данных, какова цель синтеза данных, а также оценивается вероятность утечки данных (см. ГОСТ Р 70889).

Ожидаемые результаты:

- определены цели и задачи синтеза данных;
- идентифицированы необходимые источники данных и потребители синтетических данных;
- оценены вероятности утечки данных, связанные с результатом синтеза данных (см. ПНСТ 1066—2026) и определена стратегия по ее минимизации.

Примечание — В настоящем стандарте под утечкой данных понимается раскрытие или восстановление исходных данных на основе анализа синтетических данных, моделей генерации или метаданных. Это определение не связано с классическими средствами защиты информации, такими как DLP-системы, и отражает риски, специфичные для технологий синтеза и генерации данных. Меры предотвращения утечки данных включают, в частности, применение RDP, ограничение генераций, фильтрацию выходов и контроль вероятности восстановления исходной информации.

5.2.2 Стадия 2 «Формирование требований»

Стадия 2 «Формирование требований» включает определение требований к синтетическим данным, включая их качество, формат и методы обеспечения приватности. На данной стадии необходимо уточнить бизнес-задачи, для которых собираются синтетические данные, и определить чувствительность исходных данных.

Ожидаемые результаты:

- сформулированы требования к качеству и формату синтетических данных;
- определены методы обеспечения приватности;
- согласованы бизнес-задачи и требования к уровню чувствительности данных.

5.2.3 Стадия 3 «Планирование работы с синтетическими данными»

На стадии 3 «Планирование работы с синтетическими данными» разрабатывается стратегия синтеза, оцениваются технологические ограничения, выбираются методы синтеза, определяются необходимые ресурсы и сроки выполнения.

Результаты:

- разработана схема синтеза данных;
- оценены технологические и проектные ограничения и выбраны методы синтеза;
- определены необходимые ресурсы и сроки.

Примечание — Рекомендуется применять подходы, включающие анализ вероятности утечки данных, оценку пригодности синтетических данных для дальнейшего использования, а также анализ альтернативных методов синтеза и их влияния на итоговое качество данных.

5.2.4 Стадия 4 «Сбор и агрегация эталонных данных»

Стадия 4 «Сбор и агрегация эталонных данных» включает получение исходных данных из различных источников, их агрегацию и предварительный анализ. На данной стадии необходимо подготовить эталонные данные, которые отражают статистические закономерности исходных данных. В зависимости от выбранного метода обеспечения приватности данных, включая RDP, может использоваться различная единица приватности (запись, субъект или транзакция), что позволяет обеспечить гибкость в применении методов обеспечения приватности.

Результаты:

- собраны и агрегированы данные из различных источников;
- проведена оценка полноты и целостности данных.

Примечание — При необходимости работы стадии 3 (уточнение требований) могут быть повторены, а также проведена повторная оценка рисков утечки данных и корректировка стратегии их минимизации.

5.2.5 Стадия 5 «Подготовка эталонных данных»

Стадия 5 «Подготовка эталонных данных» включает задачи по профилированию, нормализации, очистке, трансформации, кодированию и дополнительному агрегированию исходных данных с учетом требований к уровню доверия.

Результаты:

- данные очищены и трансформированы;

- выделены тренировочные и тестовые наборы данных (опционально для оценки качества синтеза и последующей валидации);

- проведено профилирование данных.

Примечание — Профилирование и очистка данных могут учитывать особенности конкретных данных, такие как наличие выбросов, пропущенных значений и других аномалий. Также рекомендуется рассмотреть возможность использования различных методов нормализации и кодирования данных, алгоритмов распознавания сущностей.

5.2.6 Стадия 6 «Подстройка параметров модели»

Стадия 6 «Подстройка параметров модели» включает определение или уточнение параметров модели МО для оптимальной работы синтеза данных.

Примечания

1 Стадия 6 может быть объединена со стадией 7 в единый обучающий цикл.

2 Описываемый синтез выполняется с помощью систем ИИ и, в частности, с помощью нейронных сетей, опирается на модели МО. Модели строятся с учетом класса моделей, архитектуры моделирования, например генеративно-состязательные сети, вариационные автокодировщики и другие архитектуры, выбранные на стадии планирования. Такие модели, кроме весов генерации, зачастую содержат дополнительные параметры: уровень шума для RDP, настройка скорости обучения и т. п. Параметры модели подстраиваются в цикле обучения модели, но как правило требуют инициализации.

Результаты:

- настроены параметры модели для оптимальной работы;

- инициализированы дополнительные параметры, такие как уровень шума и скорость обучения.

5.2.7 Стадия 7 «Обучение модели»

Стадия 7 «Обучение модели» включает процесс итерационного обучения модели с использованием эталонных данных, настройку гиперпараметров и оценку метрик производительности.

Результаты:

- обученный генератор синтетических данных;

- оцененные и скорректированные параметры модели;

- полученные качественные и количественные метрики производительности [метрики вероятности утечки данных и качества моделирования, включая статистическую близость и полезность (см. ПНСТ 1066—2026)].

Примечание — Процесс обучения модели предполагает внимательный подход к выбору гиперпараметров и архитектуры. Рекомендуется использование различных архитектур, таких как генеративно-состязательные сети, вариационные автоэнкодеры, вероятностные сети и другие методы, с учетом их особенностей и применимости к задаче синтеза данных.

5.2.8 Стадия 8 «Синтез выходных данных»

Стадия 8 «Синтез выходных данных» включает процесс генерации окончательных синтетических данных с использованием предварительно обученной модели.

Примечания

1 Допускается интеграция данной стадии с обучением модели, если это возможно в рамках выбранного подхода. Рекомендуется сохранить обученные модели для последующего использования в производственной среде.

2 Обученная модель или ее генерирующая часть (генератор синтетических данных) могут быть сохранены для последующего применения и синтеза данных в условиях эксплуатации системы ИИ.

Результаты:

- сгенерированы синтетические данные;

- проведена оценка синтетических данных.

5.2.9 Стадия 9 «Постпроцессинг синтетических данных»

Стадия 9 «Постпроцессинг синтетических данных» — это процесс завершающей обработки синтетических данных, включая их денормализацию и декодирование.

Результаты:

- данные восстановлены для использования за пределами системы ИИ;

- применены дополнительные методы фильтрации, обеспечивающие требуемый уровень доверия к технологиям синтеза и качества данных, проверены требования приватности.

5.2.10 Стадия 10 «Валидация и оценка качества синтетических данных»

Стадия 10 «Валидация и оценка качества синтетических данных» включает проверку синтетических данных на соответствие установленным требованиям.

Примечание — Данная стадия может пересекаться и частично выполняться в процессе обучения, как часть промежуточных оценок данных после выполнения очередных итераций обучающего цикла.

Результаты:

- вычислены метрики производительности синтеза;
- оценены вероятности утечки данных (см. ПНСТ 1064—2026, пункт 3.3.2);
- подтверждено отсутствие ПДн в синтетических данных;
- оценена полезность синтетических данных.

5.2.11 Стадия 11 «Публикация синтетических данных»

Стадия 11 «Публикация синтетических данных» включает размещение синтетических данных в хранилищах и обеспечение управляемого доступа к ним.

Результаты:

- синтетические данные размещены в хранилищах;
- обеспечен контролируемый доступ к синтетическим данным с учетом требований к доверенным технологиям ИИ.

5.2.12 Стадия 12 «Развертывание модели искусственного интеллекта»

Стадия 12 «Развертывание модели искусственного интеллекта» включает внедрение обученной модели в производственную среду для повторного использования.

Примечание — Стадия 12 включает в себя как разработку модели, так и имплементацию подсистем, необходимых для подготовки данных.

Результаты:

- генератор синтетических данных внедрен в производственную среду;
- обеспечена возможность повторного использования обученной модели с учетом ограничений вероятности утечки данных (опционально).

Примечания

1 Неограниченное использование модели для создания множества наборов синтетических данных может привести к нарушению уровня вероятности утечки данных. Контроль применения такого сценария осуществляется на основе таких метрик, как бюджет приватности в случае RDP (см. ПНСТ 1066—2026, пункт 5.4.3).

2 Если генератор синтетических данных разработан для определенного класса данных или задач (например, типичные финансовые транзакции), основные шаги и решения на стадиях подготовки 1—4 (конвейер обработки данных) могут использоваться повторно без необходимости полного анализа и настройки параметров. Однако при достижении лимита синтеза в рамках установленного бюджета приватности (см. В.2.5) выполняется обновление либо весов модели, либо набора данных.

5.2.13 Стадия 13 «Эксплуатация системы искусственного интеллекта»

Стадия 13 «Эксплуатация системы искусственного интеллекта» включает поддержку и мониторинг работы модели ИИ в производственной среде.

Примечание — Рекомендуется внедрение механизмов непрерывного мониторинга, оптимизации модели и регулярной переоценки вероятности утечки данных.

Результаты:

- обеспечено функционирование системы ИИ;
- проведен мониторинг и оптимизация модели.

5.2.14 Стадия 14 «Вывод синтетических данных из эксплуатации»

Стадия 14 «Вывод синтетических данных из эксплуатации» включает удаление или архивацию синтетических данных по завершении их использования.

Результаты:

- синтетические данные удалены или архивированы в соответствии с политиками обработки данных;
- в системных журналах произведены отметки об окончании поддержки синтетических данных в текущей версии, сопутствующая метаинформация удалена или архивирована для последующего анализа.

5.2.15 Стадия 15 «Вывод из эксплуатации системы»

Стадия 15 «Вывод из эксплуатации системы» включает отключение модели ИИ от производственных систем, архивирование и удаление модели.

Примечание — При выводе модели из эксплуатации следует обеспечить приватность данных, связанных с моделью.

Результаты:

- система ИИ отключена от производственных систем;
- модель удалена или архивирована.

6 Эталонная модель процесса синтеза данных

6.1 Описание процессов синтеза

Процессы синтеза данных представляют собой комплекс задач и действий, направленных на создание синтетических данных. В совокупности процессы образуют PRM, как это определено в ГОСТ Р ИСО/МЭК 33001. Модель описывает процессы с точки зрения их целей и результатов, а также взаимосвязи между ними.

Процессы синтеза данных подразделяются на три основные категории:

- поддерживающие процессы: включают управление инфраструктурой, знаниями и информацией, конфигурациями и документацией;
- управляющие процессы: включают планирование, управление уровнем доверия процесса синтеза, определение требований, обеспечение качества и систему обратной связи;
- технические процессы: охватывают сбор и подготовку исходных данных, разработку и тестирование модели, генерацию и валидацию синтетических данных, а также развертывание модели и управление уровнем доверия синтеза данных (см. [1]).

Примечание — Настоящий стандарт охватывает только процессы, напрямую относящиеся к синтезу данных. Процессы, такие как процессы соглашения, взаимодействие с заинтересованными сторонами и развитие компетенций, не рассматриваются. Для процессов этого типа рекомендуется использовать более общие модели ИТ, описанные в ГОСТ Р ИСО/МЭК 24668, ГОСТ Р 57640 и ГОСТ Р 70889.

На рисунке 1 представлена общая диаграмма процессов синтеза.

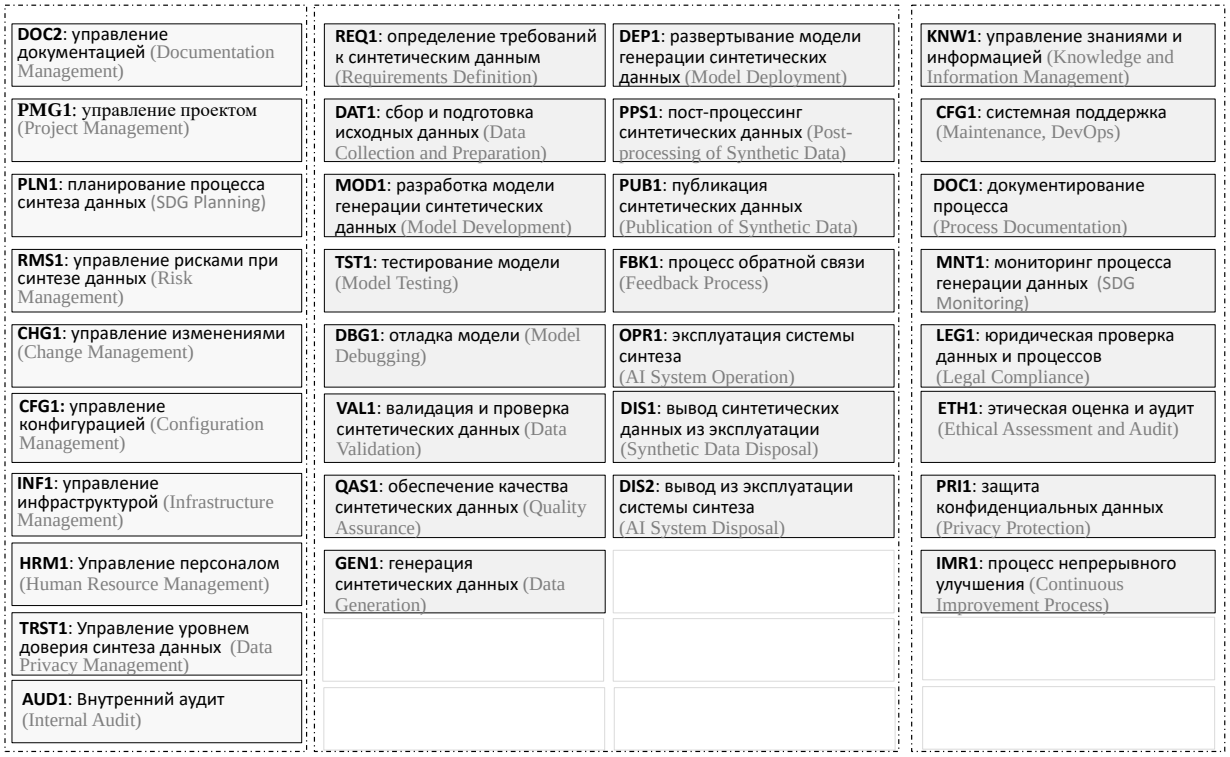


Рисунок 1 — Диаграмма процессов синтеза данных

Примечания

1 Процессы, связанные с управлением доверенными технологиями ИИ и приватностью, детализируются с учетом конкретных целей и ожидаемых результатов синтеза данных, как предписывает ГОСТ Р 57640. Эти процессы необходимы для обеспечения уровня приватности на всех стадиях их обработки.

2 Процессы, представленные в 6.2—6.4, описаны в соответствии с ГОСТ Р 57098. Эти описания включают идентификатор, наименование, цели, результаты и контекст применения для каждого подпроцесса. Описания процессов не включают характеристики качества, так как этот аспект не входит в определение процессов синтеза.

6.2 Управляющие подпроцессы

6.2.1 Управление проектом

Процесс управления проектом представлен в таблице 1.

Таблица 1 – Процесс управления проектом

Идентификатор процесса	PMG1
Название	Управление проектом
Цель	Координация всех аспектов проекта, включая ресурсы, сроки, коммуникации и выполнение задач
Контекст	Процесс является сквозным, охватывая все стадии ЖЦ синтетических данных
Выходные результаты	План проекта, отчеты о ходе выполнения, корректировки плана
Ссылка	ГОСТ Р 71539–2024 (пункт 6.3.2)

6.2.2 Планирование процесса синтеза данных

Процесс планирования представлен в таблице 2.

Таблица 2 – Процесс планирования

Идентификатор процесса	PLN1
Название	Планирование процесса синтеза данных
Цель	Определение целей, задач, ресурсов и временных рамок для процесса генерации данных
Контекст	Процесс нацелен на подготовку к синтезу и включает определение организационных параметров синтеза, охватывая стадию планирования в ЖЦ синтетических данных
Выходные результаты	План проекта, включающий цели, задачи, ресурсы и временные рамки
Ссылка	ГОСТ Р 71539–2024 (пункт 6.3.1)

6.2.3 Управление рисками при синтезе данных

Процесс управления рисками при синтезе данных представлен в таблице 3.

Таблица 3 – Процесс управления рисками при синтезе данных

Идентификатор процесса	RMS1
Название	Управление рисками при синтезе данных
Цель	Идентификация, оценка и управление вероятностями утечки данных, связанными с генерацией синтетических данных
Контекст	Идентификация, оценка и управление возможными утечками данных, связанными с синтезом данных. Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных, с особым акцентом на управление вероятностями и их снижение
Выходные результаты	План управления вероятностями утечки данных, включающий идентифицированные возможности утечки и методы их минимизации
Ссылка	ГОСТ Р 71440; ГОСТ Р 71539–2024 (пункт 6.3.4)

6.2.4 Управление изменениями

Процесс управления изменениями представлен в таблице 4.

Таблица 4 – Процесс управления изменениями

Идентификатор процесса	CHG1
Название	Управление изменениями
Цель	Обеспечение контролируемого и структурированного процесса изменений
Контекст	Процесс охватывает все стадии ЖЦ синтетических данных, обеспечивая управление изменениями на каждой стадии
Выходные результаты	Записи об изменениях, утвержденные изменения, отчеты об изменениях
Ссылка	ГОСТ Р 71539–2024 (пункты 6.3.3, 6.3.7)

6.2.5 Управление конфигурацией

Процесс управления конфигурацией представлен в таблице 5.

Таблица 5 – Процесс управления конфигурацией

Идентификатор процесса	CFG1
Название	Управление конфигурацией
Цель	Обеспечение целостности и согласованности конфигураций всех компонентов системы
Контекст	Процесс является сквозным и проходит через все стадии ЖЦ синтетических данных
Выходные результаты	Записи о конфигурациях и отчеты, подтверждающие целостность и согласованность компонентов системы
Ссылка	ГОСТ Р 71539—2024 (пункт 6.3.5)

6.2.6 Управление уровнем доверия синтеза данных

Процесс управления уровнем доверия синтеза данных представлен в таблице 6.

Таблица 6 – Процесс управления уровнем доверия синтеза данных

Идентификатор процесса	TRST1
Название	Управление уровнем доверия синтеза данных
Цель	Обеспечение приватности на всех стадиях ЖЦ синтетических данных.
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных.
Выходные результаты	Отчеты о приватности, реализованные меры по обеспечению необходимого уровня приватности.
Ссылка	Специфический процесс для синтеза

6.2.7 Управление инфраструктурой

Процесс управления инфраструктурой представлен в таблице 7.

Таблица 7 – Процесс управления инфраструктурой

Идентификатор процесса	INF1
Название	Управление инфраструктурой
Цель	Обеспечение необходимой ИТ-инфраструктуры для всех стадий ЖЦ синтетических данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Обновленная и поддерживаемая ИТ-инфраструктура, необходимая для обеспечения работы системы синтеза данных
Ссылка	ГОСТ Р 71539–2024 (пункт 6.2.2)

6.2.8 Внутренний аудит

Процесс внутреннего аудита представлен в таблице 8.

Таблица 8 – Процесс внутреннего аудита

Идентификатор процесса	AUD1
Название	Внутренний аудит
Цель	Проведение внутреннего аудита для обеспечения соответствия проектной деятельности установленным стандартам и требованиям
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Подготовленные отчеты о внутреннем аудите и рекомендации по улучшению процессов синтеза данных
Ссылка	ГОСТ Р 71539

6.2.9 Управление документацией

Процесс управления документацией представлен в таблице 9.

Таблица 9 – Процесс управления документацией

Идентификатор процесса	DOC2
Название	Управление документацией
Цель	Организация и управление документацией, связанной с процессами синтеза данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Организованная и структурированная документация, доступная для всех участников проекта синтеза
Ссылка	ГОСТ Р 71539–2024 (пункт 6.2.6)

6.2.10 Управление персоналом

Процесс управления персоналом представлен в таблице 10.

Таблица 10 – Процесс управления персоналом

Идентификатор процесса	HRM1
Название	Управление персоналом
Цель	Подбор, обучение и управление персоналом, задействованным в процессах синтеза данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Квалифицированный персонал, готовый к выполнению задач по генерации и управлению синтетическими данными
Ссылка	ГОСТ Р 71539—2024 (пункт 6.2.4)

6.3 Основные технические процессы

6.3.1 Определение требований к синтетическим данным

Таблица 11 – Процесс управления требованиями

Идентификатор процесса	REQ1
Название	Определение требований к синтетическим данным
Цель	Сбор и анализ требований от всех заинтересованных сторон, включая требования к качеству, формату и полноте реализации синтеза
Контекст	Сбор и анализ требований от всех заинтересованных сторон, касающихся синтетических данных
Выходные результаты	Документ с требованиями к синтетическим данным, включая качество, формат синтетических данных и методы обеспечения уровня приватности
Ссылка	ГОСТ Р 57193, ГОСТ Р 71539—2024 (пункт 6.4.3)

6.3.2 Сбор и подготовка исходных данных

Процесс сбора и подготовки исходных данных представлен в таблице 12.

Таблица 12 – Процесс сбора и подготовки исходных данных

Идентификатор процесса	DAT1
Название	Сбор и подготовка исходных данных
Цель	Сбор, очистка и предобработка реальных данных для использования в процессе генерации синтетических данных, включая описание происхождения и меры по обеспечению приватности
Контекст	Процесс охватывает стадии сбора и подготовки данных в ЖЦ синтетических данных
Выходные результаты	Подготовленные и очищенные данные, готовые для использования в моделях генерации полных синтетических данных
Ссылка	ГОСТ Р 71539–2024 (пункт 6.4.7)

6.3.3 Разработка модели генерации синтетических данных

Процесс разработки модели генерации синтетических данных представлен в таблице 13.

Таблица 13 – Процесс разработки модели генерации синтетических данных

Идентификатор процесса	MOD1
Название	Разработка модели генерации синтетических данных
Цель	Проектирование и разработка модели для генерации синтетических данных с учетом требований по уровню доверия к технологиям синтеза
Контекст	Процесс охватывает стадии разработки модели в ЖЦ синтетических данных (обучение и подстройка параметров модели)
Выходные результаты	Разработанная и протестированная модель генерации синтетических данных, готовая для использования
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.5)

6.3.4 Тестирование модели

Процесс тестирования модели представлен в таблице 14.

Таблица 14 – Процесс тестирования модели

Идентификатор процесса	TST1
Название	Тестирование модели
Цель	Обеспечение корректной работы модели, соответствия требованиям и гарантиям приватности данных
Контекст	Процесс проходит на стадии валидации и оценки качества в ЖЦ синтетических данных
Выходные результаты	Результаты тестирования, подтверждающие работоспособность, качество и соблюдение требований к доверенным технологиям ИИ
Ссылка	ГОСТ Р 71539

6.3.5 Отладка модели

Процесс отладки модели представлен в таблице 15.

Таблица 15 – Процесс отладки модели

Идентификатор процесса	DBG1
Название	Отладка модели
Цель	Исправление ошибок и проблем, выявленных на стадии тестирования модели
Контекст	Процесс включается в стадии обучения модели и подстройки параметров (цикл обучения) в ЖЦ синтетических данных
Выходные результаты	Исправленная, оптимизированная модель генерации синтетических данных с заданным уровнем приватности
Ссылка	ГОСТ Р 71539

6.3.6 Валидация синтетических данных

Процесс валидации синтетических данных представлен в таблице 16.

Таблица 16 – Процесс валидации синтетических данных

Идентификатор процесса	VAL1
Название	Валидация синтетических данных
Цель	Проверка синтетических данных на соответствие требованиям качества и уровням доверия процесса синтеза
Контекст	Процесс охватывает стадию валидации синтетических данных в ЖЦ
Выходные результаты	Отчеты о проверке синтетических данных, подтверждающие их соответствие требованиям качества и низким значениям вероятности утечки данных
Ссылка	ГОСТ Р 71539

6.3.7 Обеспечение качества синтетических данных

Процесс обеспечения качества синтетических данных представлен в таблице 17.

Таблица 17 – Процесс обеспечения качества синтетических данных

Идентификатор процесса	QAS1
Название	Обеспечение качества синтетических данных
Цель	Обеспечение соответствия синтетических данных требованиям, стандартам качества и уровню доверия к технологиям синтеза
Контекст	Обеспечение соответствия синтетических данных требованиям и стандартам качества. Процесс является сквозным
Выходные результаты	Отчеты о качестве, включающие результаты проверок, тестов (в том числе отсутствия ПДн в синтетических данных) и оценки вероятности утечки данных
Ссылка	ГОСТ Р 57193—2025 (пункт 6.4.11), ГОСТ Р 71539–2024 (пункт 6.3.8)

6.3.8 Генерация синтетических данных

Процесс генерации синтетических данных представлен в таблице 18.

Таблица 18 – Процесс генерации синтетических данных

Идентификатор процесса	GEN1
Название	Генерация синтетических данных
Цель	Использование обученной модели для создания синтетических данных
Контекст	Процесс охватывает стадию генерации синтетических данных в ЖЦ
Выходные результаты	Синтетические данные, соответствующие требованиям качества и уровню доверия
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.9)

6.3.9 Развертывание модели генерации

Процесс развертывания модели генерации представлен в таблице 19.

Таблица 19 – Процесс развертывания модели генерации

Идентификатор процесса	DEP1
Название	Развертывание модели генерации
Цель	Интеграция и развертывание модели генерации синтетических данных в эксплуатационную среду с соблюдением требований к уровню доверия технологий синтеза
Контекст	Процесс охватывает стадию развертывания модели в ЖЦ синтетических данных
Выходные результаты	Генератор синтетических данных, развернутый в производственной среде, готовый к использованию и обеспечивающий низкий уровень вероятности утечки данных
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.12)

6.3.10 Постпроцессинг синтетических данных

Процесс пост-обработки синтетических данных представлен в таблице 20.

Таблица 20 – Процесс пост-обработки синтетических данных

Идентификатор процесса	PPS1
Название	Постпроцессинг синтетических данных
Цель	Завершающая обработка синтетических данных, включая денормализацию, декодирование, проверку приватности и возможности утечки данных
Контекст	Процесс охватывает стадию пост-обработки в ЖЦ синтетических данных
Выходные результаты	Обработанные синтетические данные, готовые для использования
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.12)

6.3.11 Публикация синтетических данных

Процесс публикации синтетических данных представлен в таблице 21.

Таблица 21 – Процесс публикации синтетических данных

Идентификатор процесса	PUB1
Название	Публикация синтетических данных
Цель	Размещение синтетических данных в общедоступных или закрытых хранилищах с учетом требований к полностью синтетическим данным для использования
Контекст	Процесс охватывает стадию публикации в ЖЦ синтетических данных
Выходные результаты	Опубликованные синтетические данные, доступные для пользователей
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.10)

6.3.12 Процесс обратной связи

Процесс поддержки обратной связи представлен в таблице 22.

Таблица 22 – Процесс поддержки обратной связи

Идентификатор процесса	FBK1
Название	Процесс обратной связи
Цель	Сбор и анализ обратной связи от пользователей синтетических данных для повышения качества и уровня доверия к использованным технологиям
Контекст	Сбор и анализ обратной связи от пользователей синтетических данных. Процесс относится к стадии публикации синтетических данных
Выходные результаты	Отчеты с обратной связью, включающие рекомендации по улучшению процесса генерации данных
Ссылка	ГОСТ Р 57193, ГОСТ Р 71539—2024 (пункт 6.4.2)

6.3.13 Эксплуатация системы синтеза

Процесс эксплуатации системы синтеза представлен в таблице 23.

Таблица 23 – Процесс эксплуатации системы синтеза

Идентификатор процесса	OPR1
Название	Эксплуатация системы синтеза
Цель	Поддержка и мониторинг работы модели синтеза в производственной среде с акцентом на доверенные технологии ИИ
Контекст	Процесс охватывает стадию эксплуатации в ЖЦ синтетических данных
Выходные результаты	Рабочая система синтеза, оптимизированная и обновленная по мере необходимости
Ссылка	ГОСТ Р 57193, ГОСТ Р 71539—2024 (пункт 6.4.15)

6.3.14 Вывод синтетических данных из эксплуатации

Процесс вывода синтетических данных из эксплуатации представлен в таблице 24.

Таблица 24 – Процесс вывода синтетических данных из эксплуатации

Идентификатор процесса	DIS1
Название	Вывод синтетических данных из эксплуатации
Цель	Удаление или архивация синтетических данных по завершении их использования
Контекст	Процесс охватывает стадию вывода данных из эксплуатации в ЖЦ синтетических данных
Выходные результаты	Удаленные или архивированные синтетические данные
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.17)

6.3.15 Вывод из эксплуатации системы синтеза

Вывод из эксплуатации системы синтеза представлен в таблице 25.

Таблица 25 – Процесс вывода из эксплуатации системы синтеза

Идентификатор процесса	DIS2
Название	Вывод из эксплуатации системы синтеза
Цель	Отключение модели ИИ от производственных систем, удаление или архивирование модели
Контекст	Процесс охватывает стадию вывода из эксплуатации в ЖЦ синтетических данных
Выходные результаты	Отключенная и архивированная модель ИИ (генератор синтетических данных)
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.17)

6.4 Поддерживающие процессы

6.4.1 Управление знаниями и информацией

Процесс управления знаниями и информацией представлен в таблице 26.

Таблица 26 – Процесс управления знаниями и информацией

Идентификатор процесса	KNW1
Название	Управление знаниями и информацией
Цель	Систематизация и управление знаниями и информацией, критически важной для синтеза данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Обновленная и структурированная база знаний и информации, доступная для участников проекта
Ссылка	ГОСТ Р 57193—2025 (пункт 5.6.3), ГОСТ Р 71539–2024 (пункт 6.2.6)

6.4.2 Системная поддержка

Процесс системной поддержки представлен в таблице 27.

Таблица 27 – Процесс системной поддержки

Идентификатор процесса	CFG1
Название	Системная поддержка
Цель	Обеспечение целостности и согласованности конфигураций всех компонентов системы синтеза данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Конфигурационные записи, отчеты о конфигурации
Ссылка	ГОСТ Р 57193—2025 (подраздел 6.2), ГОСТ Р 71539—2024 (пункт 6.4.16)

6.4.3 Документирование процесса

Процесс документирования представлен в таблице 28.

Таблица 28 – Процесс документирования

Идентификатор процесса	DOC1
Название	Документирование процесса
Цель	Создание и поддержка документации всех стадий процесса синтеза данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Полная и актуальная документация всех стадий процесса генерации синтетических данных, включая меры обеспечения доверенных технологий ИИ, и основанных на таких технологиях синтеза данных
Ссылка	ГОСТ Р 71539—2024 (пункт 6.3.6)

6.4.4 Мониторинг процесса генерации данных

Мониторинг процесса генерации данных представлен в таблице 29.

Таблица 29 – Мониторинг процесса генерации данных

Идентификатор процесса	MNT1
Название	Мониторинг процесса генерации данных
Цель	Постоянный мониторинг процесса генерации данных для обеспечения качества и создания полностью синтетических данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Отчеты о мониторинге, выявленные проблемы и меры по их устранению
Ссылка	ГОСТ Р 71539—2024 (пункт 6.4.14)

6.4.5 Юридическая проверка данных и процессов

Процесс юридической проверки представлен в таблице 30.

Таблица 30 – Процесс юридического сопровождения

Идентификатор процесса	LEG1
Название	Юридическая проверка данных и процессов
Цель	Обеспечение соответствия правовым нормам и стандартам
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Отчеты о юридической проверке данных и процессов, рекомендации по устранению несоответствий
Ссылка	—

6.4.6 Оценка смещенности

Процесс оценки смещенности представлен в таблице 31.

Таблица 31 – Процесс оценки смещенности

Идентификатор процесса	ETH1
Название	Процесс оценки смещенности
Цель	Оценка смещенности, моделей и соблюдения уровня доверия к использованным технологиям синтеза

Окончание таблицы 31

Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Отчеты об оценке смещенности, рекомендации по улучшению
Ссылка	—

6.4.7 Безопасность и защита персональных данных

Процесс обеспечения безопасности и защиты ПДн представлен в таблице 32.

Таблица 32 – Процесс обеспечения уровня приватности

Идентификатор процесса	PRI1
Название	Безопасность и защита персональных данных
Цель	Реализация мер по предотвращению утечек данных на всех стадиях ЖЦ синтетических данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Отчеты об уровне приватности данных, отчет о реализованных мерах защиты технологий ИИ
Ссылка	ГОСТ Р 71539–2024 (пункт 6.4.14)

6.4.8 Процесс непрерывного улучшения

Процесс непрерывного улучшения представлен в таблице 33.

Таблица 33 – Процесс непрерывного улучшения

Идентификатор процесса	IMR1
Название	Процесс непрерывного улучшения
Цель	Внедрение постоянных улучшений в процессы генерации, использования и снижения вероятности утечки данных на основании синтетических данных
Контекст	Процесс является сквозным и охватывает все стадии ЖЦ синтетических данных
Выходные результаты	Рекомендации по улучшению, реализованные улучшения
Ссылка	ГОСТ Р ИСО 9000, ГОСТ Р 71539—2024 (пункт 6.4.14)

7 Эталонная архитектура системы синтеза данных

7.1 Общие положения

Эталонная архитектура представляет собой концептуальную структуру, описывающую ключевые компоненты, их взаимосвязи, а также правила и ограничения, характерные для всех систем в определенной предметной области. В данном разделе рассматривается эталонная архитектура для систем синтеза данных на основе ИИ. Архитектура системы включает описание структуры, взаимосвязей компонентов, а также правил и ограничений, что обеспечивает целостное и системное представление функционирования системы (см. [2]). Проектирование архитектуры должно учитывать гарантии приватности и критерии качества системы ИИ, установленные в ГОСТ Р 59898.

Структура архитектуры может быть представлена в виде совокупности следующих компонентов:

- модули — функциональные единицы, реализующие отдельные задачи или операции внутри системы. Взаимодействие между модулями позволяет организовать выполнение составных процессов в рамках синтеза данных;

- подсистемы — это логически объединенные группы модулей, ориентированные на решение более общих задач. Они формируются по функциональному или технологическому признаку в зависимости от особенностей архитектуры;

- слои — уровни организации компонентов системы, применяемые для упрощения проектирования, управления и взаимодействия. Слои объединяют соответствующие подсистемы и модули, отражая их роль в архитектуре.

К основным заинтересованным лицам для систем синтеза относятся поставщики инфраструктуры данных, пользователи синтетических данных, разработчики систем ИИ, инженеры по данным, специализированные подразделения, владельцы исходных данных и регуляторы. Роли и действия заинтересованных лиц должны быть учтены при реализации системы синтеза данных. Рекомендации к проектированию системы синтеза данных изложены в приложении Б. Технические особенности и методы синтеза приведены в приложении В.

Примечания

1 Настоящий стандарт определяет архитектуру систем синтеза данных, расширяя эталонную архитектуру для больших данных и систем ИИ в соответствии с ГОСТ Р 70466 и перспективными направлениями стандартизации.

2 При использовании федеративных систем МО для синтеза данных архитектура может проектироваться с учетом положений, установленных для распределенных систем.

3 При проектировании архитектуры могут учитываться критерии качества системы ИИ, установленные в ГОСТ Р 59898.

7.2 Функциональная архитектура синтеза данных

7.2.1 Состав и назначение функциональной архитектуры

Функциональная архитектура синтеза данных представляет собой структурированное описание компонентов и их взаимодействий, обеспечивающих эффективное и надежное создание полностью синтетических данных (см. [1]). Основное предназначение данной архитектуры — обеспечить целостное и систематическое представление всех элементов и процессов, участвующих в синтезе данных, что позволяет повысить управляемость, масштабируемость и надежность системы.

Эталонную функциональную архитектуру следует применять во всех случаях, где требуется синтез данных с использованием ИИ, при проектировании систем промышленного синтеза данных.

При реализации функциональной архитектуры должны выполняться следующие требования:

- архитектура должна учитывать специфику обработки больших данных, интеграции систем ИИ и обеспечения приватности данных на всех стадиях ЖЦ синтеза данных;

- все компоненты должны соответствовать требованиям ГОСТ Р ИСО/МЭК 20547-3;

- при проектировании подсистем обеспечения уровня доверия должны выполняться требования ГОСТ Р 59347 (в процессе определения архитектуры системы).

Функциональная архитектура должна включать следующие слои:

- слой данных — покрывает функции хранения данных в системе, как исходных, так и синтетических, полученных в результате синтеза, управление метаинформацией и сопутствующими метриками;

- слой обработки данных — включает все функции обработки данных, например, такие как подготовку исходных данных, обучение моделей и др., а также постпроцессинг и подготовку результата синтеза;

- слой интеграции — включает функции интеграции системы синтеза с внешними системами, а также обеспечение механизмов публикации данных и контроля доступа;

- слой приватности — объединяет подсистемы обеспечения уровня доверия к использованным технологиям синтеза.

Примечание – Слой поддерживает специфические системы предотвращения утечек данных, такие как RDP или фильтрация выходных данных, гарантирующие приватность получаемых синтетических данных (полных синтетических данных). При проектировании принимают во внимание особенности организации контуров доступа, отделяющих чувствительные данные от получаемых промышленных данных;

- слой управления системой — поддерживает подсистемы контроля вероятности утечки данных, непрерывного управления качеством и мониторингом.

7.2.2 Слой данных

Слой данных отвечает за надежное и эффективное хранение всех данных, связанных с процессом синтеза данных. Основные функции слоя включают:

- хранение исходных данных: обеспечивает хранение эталонных данных для обучения и тестирования моделей;

- хранение синтетических данных: обеспечивает сохранение синтетических данных с соблюдением требований к уровню доверия технологий синтеза;

- хранение моделей генерации: обеспечивает сохранение и управление версиями моделей генерации данных.

Компоненты слоя данных:

- входные данные: подсистема «Входные данные» отвечает за сбор, хранение и накопление исходных данных;

- хранилище синтетических данных и публикация: обеспечивает хранение и публикацию синтетических данных;

- хранилище моделей: управляет ЖЦ моделей, включая их обновление и контроль версий.

Основные требования к подсистемам слоя данных:

- данные должны храниться в соответствии с требованиями приватности, учитывая установленные для организации контуры доступа;

- компоненты слоя данных должны соответствовать требованиям к инфраструктурным уровням, описанным в ГОСТ Р ИСО/МЭК 20547-3.

7.2.3 Слой обработки данных

Слой обработки данных выполняет функции предобработки данных, их синтеза, постобработки и валидации.

Обязательные компоненты и функции:

- предобработка данных: может включать нормализацию, профилирование, очистку, агрегацию и трансформацию данных;

- обучение и настройка модели: отвечает за разработку, настройку и обучение моделей для генерации синтетических данных;

- постобработка данных: отвечает за обработку синтетических данных после процесса синтеза для их последующего использования;

- валидация и оценка качества: оценка синтетических данных на соответствие требованиям качества и приватности. При оценке качества синтетических данных должны учитываться показатели статистической близости, полезности и уровня приватности, а также возможное влияние повышения статистической близости на вероятность утечки исходной информации.

При реализации слоя обработки данных системы синтеза следует руководствоваться следующими требованиями:

- подсистемы слоя должны функционировать в закрытом контуре, обеспечивая отсутствие утечек данных на всех стадиях их обработки;

- все компоненты слоя обработки данных должны соответствовать стандартам, описанным в ГОСТ Р ИСО/МЭК 20547-3.

7.2.4 Слой интеграции

Слой интеграции представляет интерфейсы для конечных потребителей данных и обеспечивает возможность повторного использования обученных моделей.

Компоненты слоя интеграции:

- вывод данных: обеспечивает надежную генерацию полностью синтетических данных для производственного использования, развертывание модели и контроль доступа;

- интерфейсы и интеграция: обеспечивает внешние интерфейсы и модули для интеграции с другими системами.

Все интеграционные компоненты должны обеспечивать совместимость с внешними системами и соответствовать стандартам интерфейсов, установленным ГОСТ Р ИСО/МЭК 24668.

7.2.5 Слой приватности

Слой приватности (Privacy Layer) включает функции обеспечения качественной реализации доверенных технологий ИИ и предотвращение утечек данных через синтетические данные.

Компоненты слоя приватности:

- механизмы реализации доверенных технологий ИИ: включают RDP, фильтрацию данных и анонимизацию;
- обеспечение доверия: включает аудиторские модули и методы объяснимости моделей для поддержания доверия к системе.

Слой приватности должен учитывать требования по контролю вероятности утечки данных, управлению потенциальными проблемами и соблюдению установленных стандартов обработки данных:

- все системы контроля вероятности утечки данных должны быть размещены внутри закрытого контура и обеспечивать надежную анонимизацию и снижение вероятности утечек данных;
- слой приватности должен включать управление показателями для измерения вероятности утечки данных при публикации с учетом сценария взаимодействия;
- механизмы обеспечения уровня доверия к технологиям синтеза должны соответствовать ГОСТ Р 71440 и ГОСТ Р 59347.

7.2.6 Слой управления системой

Слой управления системой (System Management Layer) содержит модули для управления процессами синтеза данных и обеспечения их непрерывной работы.

Компоненты слоя управления системой:

- управление процессом синтеза и планирование операций: включает планирование и управление процессами синтеза данных;
- управление задачами генерации: отвечает за запуск задач генерации и управление их статусом;

- мониторинг производительности и логирование: обеспечивает отслеживание производительности системы и логирование операций;

- конфигурация и управление версиями: управляет конфигурациями системы и версиями компонентов;

- непрерывное управление качеством работы: включает сбор метрик, обеспечение согласованности процессов, обработку журналов и обеспечение соответствия процесса заданным целям;

- постоянное управление уровнем доверия: включает идентификацию, оценку и мониторинг вероятности утечки данных, связанных с синтезом данных.

При реализации слоя управления должны учитываться следующие требования:

- процессы управления должны быть интегрированы с системой мониторинга и управления уровнем доверия, обеспечивая своевременное выявление и устранение проблем;

- компоненты управления должны соответствовать ГОСТ Р 70889.

7.3 Архитектура приватности синтеза данных

7.3.1 Приватность в системах синтеза

Типовая архитектура приватности в системах синтеза данных направлена на обеспечение уровня приватности данных на всех стадиях их обработки и хранения (см. [3]).

Системы синтеза данных представляют собой сложные распределенные информационные системы, включающие многослойные компоненты и процессы. Такие системы подвержены различным проблемам нарушения приватности, которые необходимо учитывать при проектировании и эксплуатации.

Системы синтеза данных часто обрабатывают исходные чувствительные данные, что требует принятия специальных мер по снижению вероятности утечек данных на всех уровнях архитектуры. Синтетические данные, несмотря на их искусственный характер, могут сохранять остаточные вероятности восстановления исходных данных через их статистические особенности. В связи с этим обязательным является использование методов обеспечения доверия, таких как анонимизация, фильтрация и RDP, для минимизации вероятности утечек данных.

Примечание — Методы снижения вероятности утечки данных позволяют учитывать специфику обрабатываемой информации, типы данных, подвергающихся синтезу, а также проблемы нарушения приватности, с которыми может столкнуться система. Необходимо адаптировать методы применения доверенных технологий ИИ и основанных на них методах синтеза данных в зависимости от уровня чувствительности данных и особенностей системы.

7.3.2 Требования к архитектуре приватности

Архитектура приватности синтеза данных должна включать элементы, приведенные в 7.3.2.1—7.3.2.3 (см. [3]).

7.3.2.1 Контуры доступа

Архитектура приватности систем синтеза должна предусматривать гибкую настройку контуров доступа в зависимости от уровня чувствительности обрабатываемых данных и сценариев использования. В типичной реализации рекомендуется использовать следующие контуры:

- внутренний контур: обрабатывает наиболее чувствительные (исходные) данные и защищает от инсайдеров;
- партнерский контур: ограниченный доступ к данным для партнеров с минимизацией вероятностей утечек данных;
- публичный контур: меры против внешних злоумышленников при обеспечении доступа к данным и интерфейсам.

Примечание — В зависимости от специфики системы могут быть добавлены дополнительные контуры или изменена конфигурация существующих.

7.3.2.2 Методы повышения уровня доверия

Архитектура приватности систем синтеза должна предусматривать использование различных методов обработки данных в зависимости от вероятностей утечек данных и требований системы. К таким методам относятся:

- обезличивание включает удаление или маскировку идентифицирующей информации с использованием методов k-анонимности, l-разнообразия, t-близости и других подходов, обеспечивающих снижение вероятности утечки данных;
- RDP: добавление математически контролируемого шума к данным для обеспечения приватности с учетом заранее определенного бюджета приватности;

- фильтрация данных: удаление или маскирование данных, которые могут привести к утечке данных, включая выбросы и уникальные записи;

- шифрование: применение шифрования данных при реализации хранения и передачи с использованием стандартных алгоритмов и управления ключами (например, ГОСТ Р 34.12);

- контроль доступа: управление доступом к данным на основе ролей и многофакторной аутентификации.

Примечание — Методы повышения уровня доверия выбираются и комбинируются в зависимости от конкретных выявленных проблем и требований к системе. Применение всех методов одновременно может быть избыточным; выбор определяется уровнем доверия и спецификой данных на основе оценки вероятностей утечки данных.

7.3.2.3 Процессы управления приватностью

Для обеспечения необходимого уровня приватности на всех стадиях синтеза данных должны быть реализованы следующие процессы:

- регулярный аудит и мониторинг приватности: постоянный контроль за состоянием системы, оценка вероятностей утечки данных, выявление и устранение проблем, влияющих на такую утечку;

- оценка вероятностей утечки данных и управление рисками: регулярная оценка вероятности утечек данных на всех стадиях синтеза, с расчетом метрик вероятности утечек данных, эффективности методов снижения вероятности утечки данных и соблюдения бюджета приватности;

- расчет метрик приватности: оценка и мониторинг ключевых показателей уровня доверия к использованным технологиям синтеза, включая устойчивость данных к восстановлению и вероятность утечек данных.

Примечание — Рекомендации по проектированию и реализации архитектуры приватности, а также дополнительные методы снижения вероятности утечки данных изложены в Б.3.

7.4 Архитектура интеграции синтеза данных

7.4.1 Принципы интеграции систем синтеза данных

Типовая архитектура интеграции должна обеспечивать эффективное и безопасное взаимодействие между системой синтеза данных и другими информационными системами, включая разработку и внедрение механизмов для надежного и совместимого обмена данными между внутренними компонентами системы.

Архитектура интеграции синтеза данных должна соответствовать следующим обязательным принципам:

- совместимость: использование стандартных протоколов и интерфейсов для облегчения интеграции между системами и обеспечения их совместной работы;

- приватность и доверие: защита данных на всех стадиях с помощью шифрования, контроля доступа, аутентификации и авторизации, обеспечение приемлемых показателей вероятности утечки данных на всех стадиях интеграции;

- мониторинг и аудит: внедрение постоянного контроля всех интеграционных процессов с отслеживанием событий и выявлением сбоев;

- гибкость и масштабируемость: обеспечение возможности расширения и адаптации архитектуры без потери производительности;

- документирование: фиксация всех процессов и передаваемых данных, чтобы обеспечить прозрачность, целостность и повторяемость;

- тестирование: проверка интеграционных решений на совместимость и отсутствие утечек данных до ввода в эксплуатацию.

7.4.2 Компоненты архитектуры интеграции

Архитектура интеграции системы синтеза данных направлена на обеспечение совместимости, приватности и гибкости при взаимодействии с внешними системами. Основные компоненты архитектуры должны поддерживать имеющиеся стандарты и требования к обработке данных, обеспечивая надежную и эффективную интеграцию. Список рекомендуемых компонентов отражен в таблице 34.

Таблица 34 – Основные компоненты архитектуры интеграции

Компонент	Описание	Область применения
Интерфейсы	Стандартизированные интерфейсы для взаимодействия с внешними системами и компонентами. Поддерживают протоколы, такие как REST и GraphQL, и форматы данных, такие как JSON и XML. Обеспечивают управляемый доступ к данным и сервисам системы	Взаимодействие с внешними системами, запрос и передача данных
Модули интеграции	Модули для взаимодействия с различными внешними системами, обеспечивающие конвертацию данных, управление форматами и надежную передачу данных. Поддерживают согласованность данных при передаче между системами	Обеспечение совместимости данных при их передаче между системами
Шлюзы данных	Управление потоками данных между различными контурами безопасности, включая фильтрацию, маршрутизацию и трансформацию данных. Обеспечивают исключение утечки данных в процессе передачи и мониторинг для обеспечения приватности	Обработка и маршрутизация данных между различными системами
Буферы и очереди сообщений	Обеспечивают асинхронное взаимодействие между компонентами системы, сглаживая нагрузки и обеспечивая надежность передачи данных. Очереди сообщений защищены шифрованием и поддерживают мониторинг состояния	Асинхронная передача данных, снижение нагрузки на систему

Окончание таблицы 34

Компонент	Описание	Область применения
Модули управления метаинформацией	Управление регистрацией, трассировкой метаданных. Обеспечивают целостность и прозрачность данных в процессе интеграции и синхронизацию метаданных между системами	Управление метаданными и их синхронизация между системами
Модули управления непрерывностью процессов	Обеспечивают отказоустойчивость, резервное копирование и восстановление данных, а также мониторинг и тестирование процессов восстановления после сбоев	Управление отказоустойчивостью, резервное копирование и восстановление данных

Приложение А
(рекомендуемое)
Пояснения к жизненному циклу синтеза данных

В настоящем приложении представлена рекомендуемая схема организации ЖЦ синтетических данных.

А.1 Принципиальная схема синтеза данных

Синтез данных может осуществляться различными методами, однако данный стандарт акцентирует внимание на методах, использующих системы ИИ. Эти системы обучаются на основе статистических закономерностей исходных данных для создания синтетических данных. На рисунке А.1 представлена схема ключевых стадий процесса синтеза данных.

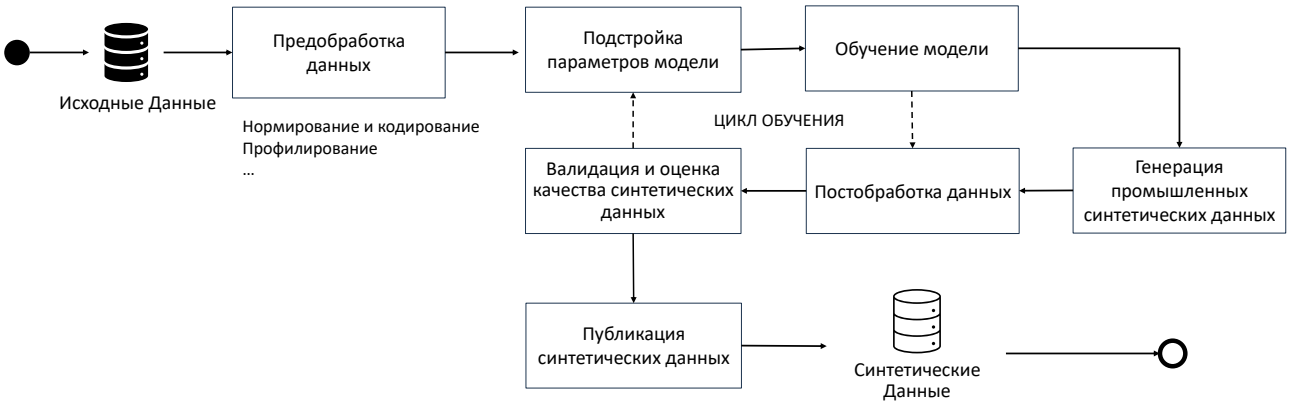


Рисунок А.1 — Принципиальная схема синтеза данных

Примечания

1 Системы ИИ проходят стадии ЖЦ, общие для любой информационной системы, согласно ГОСТ Р ИСО/МЭК 12207. Особенности ЖЦ для систем ИИ, связанных с данными, отражают взаимодействие ЖЦ системы и ЖЦ данных (синтетических данных), поддерживающих ИИ.

2 В ГОСТ Р 70889 описаны стадии обработки данных на всех стадиях ЖЦ системы ИИ, включая сбор, создание, разработку, развертывание, обслуживание и вывод из эксплуатации. Синтез данных представляет частный случай этих процессов, с акцентом на генеративный ИИ.

А.2 Основные сущности синтеза данных

Основные сущности процесса синтеза данных:

- система синтеза данных: информационная система, включающая инфраструктуру генерации данных, подсистемы сбора метрик и обеспечения уровня доверия к используемым технологиям синтеза. Система синтеза работает с исходными данными, создавая синтетические данные на основе статистических моделей;

- генератор синтетических данных (модель синтеза): математическая или компьютерная модель, отражающая статистические свойства исходных данных и позволяющая синтезировать новые данные;

- синтетические данные: данные, искусственно созданные для имитации формата и свойств реальных данных, но которые не соответствуют напрямую каким-либо реальным объектам и не являются модификацией исходных данных.

На рисунке А.2 представлена схема, показывающая взаимодействие этих сущностей в процессе синтеза данных.

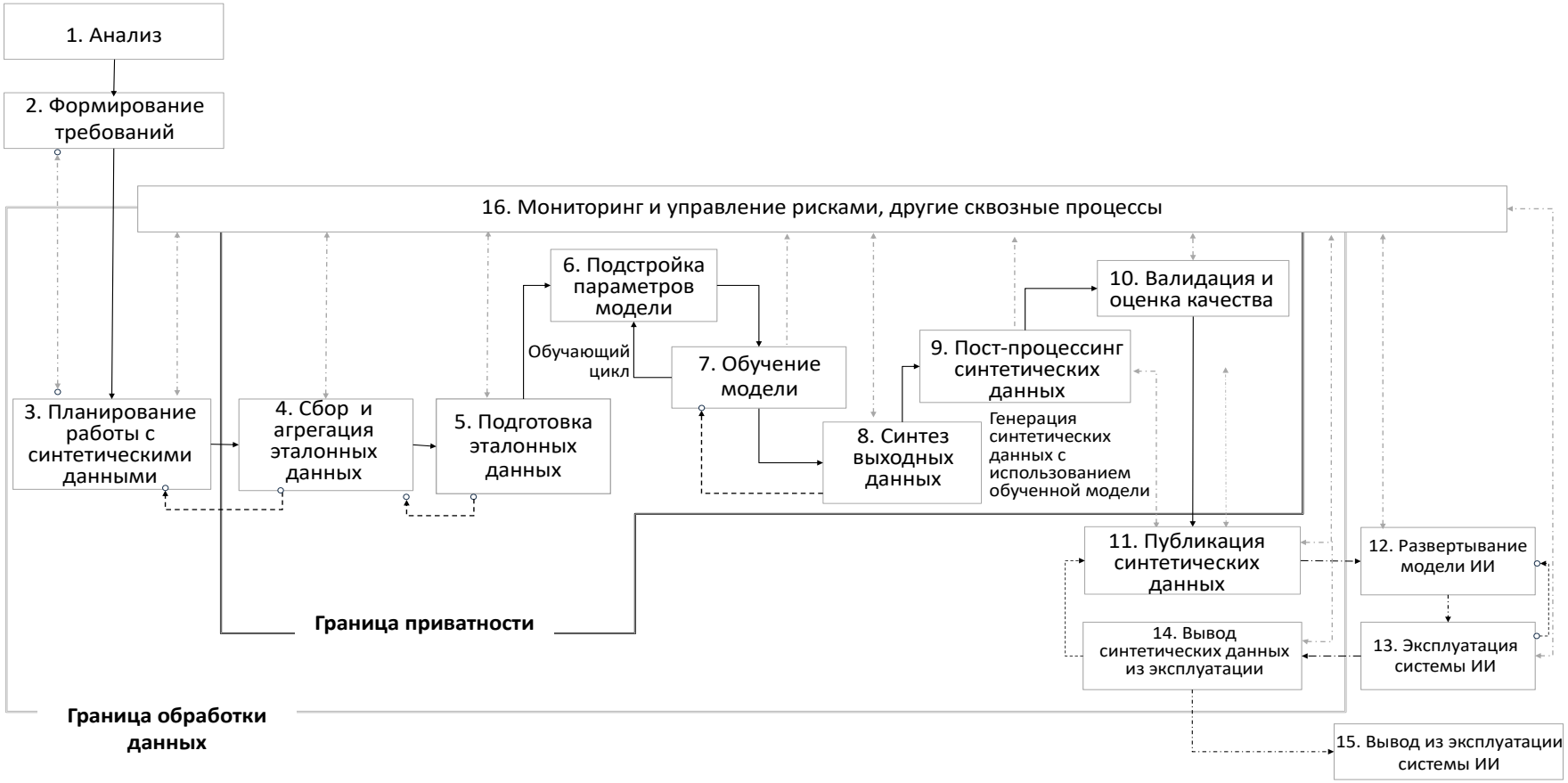
Рисунок А.2 – Основные сущности синтеза данных

А.3 Подробная структура жизненного цикла систем синтеза

ЖЦ системы синтеза данных включает несколько стадий, которые представлены на рисунке А.3. Схема ЖЦ охватывает весь процесс создания синтетических данных — от первоначального замысла и планирования до эксплуатации и вывода из обращения как синтетических данных, так и системы в целом.

Полный ЖЦ включает ключевые стадии, представленные в пункте 5.2.

В процессе осуществлений стадий ЖЦ основные сущности изменяют свои атрибуты приватности, приведенная ниже схема включает обозначение границ обработки данных и границ приватности, что особенно важно при работе с ПДн и другой информацией, на основе которой готовятся полные синтетические данные.



Примечание — На схеме отражены обратные связи между отдельными стадиями, в целом ЖЦ охватывает создание синтетических данных от первоначального замысла и планирования (стадии 1–3) до эксплуатации и вывода из обращения синтетических данных и системы в целом (стадии 12–15). Сквозные процессы (мониторинга и управления приватностью) обозначены на данной схеме. Промежуточные стадии охватывают процессы синтеза и его оценки (стадии 4–11).

Рисунок А.3 — ЖЦ системы синтеза данных

Приложение Б
(рекомендуемое)
Пояснение к архитектуре системы синтеза

В настоящем приложении представлены рекомендации в части проектирования архитектуры системы синтеза.

Б.1 Подходы к проектированию архитектуры системы синтеза

Синтез данных с использованием ИИ включает множество архитектурных решений и подходов. Описание архитектуры таких систем требует значительного уровня обобщения, допускающего уточнения в зависимости от используемого сценария.

Эталонная архитектура синтеза данных на основе ИИ основывается на ГОСТ Р 57100, который определяет принципы и методологию описания архитектуры. Основное предназначение эталонной архитектуры заключается в сосредоточении на перспективных решениях и использовании ее в качестве основы для будущих реализаций, что позволяет создавать гибкие, масштабируемые и надежные системы.

Архитектура системы синтеза данных должна учитывать свойства процессов, связанных с ЖЦ данных и систем, обеспечивая приватность, масштабируемость и гибкость. Эти свойства формируют атрибуты архитектуры, определяя ее компоненты и взаимосвязи между ними.

В соответствии с ГОСТ Р ИСО/МЭК 20547-3, архитектура подразделяется на пользовательские представления и функциональные представления, что помогает заинтересованным сторонам получить ясное понимание системы. В данном стандарте рассматриваются представления: функциональная архитектура, архитектура приватности и архитектура интеграции (см. рисунок Б.1).

Примечание — Перечисленные представления описывают эталонную реализацию и являются рекомендательными. Окончательные решения по структуре и взаимодействию компонентов принимаются с учетом целей синтеза и технологического ландшафта владельца системы.

Заинтересованные лица включают в себя разработчиков, пользователей, администраторов и регуляторов. К основным заинтересованным лицам для систем синтеза данных возможно отнести:

- владельцев исходных данных;
- поставщиков инфраструктуры данных;
- пользователей синтетических данных;
- разработчиков систем ИИ;
- инженеров безопасности и специализированные подразделения;
- поставщиков данных;
- владельцев исходных данных (например, ФЛ, если система ИИ обучает модель ИИ с использованием ПДн);

- регуляторов и представителей контролирующих органов.

Диаграмма архитектуры системы синтеза данных, показывающая пользовательские и функциональные представления, включая роли и действия заинтересованных лиц и технические аспекты системы, представлена на рисунке Б.1.

Рисунок Б.1 – Заинтересованные лица при построении архитектуры системы синтеза данных

Б.2 Проектирование функциональной архитектуры системы синтеза

Функциональная архитектура системы синтеза представляет собой набор подсистем, каждая из которых играет отдельную роль в процессе синтеза данных. Функциональная архитектура охватывает весь ЖЦ данных: от предобработки до их использования и интеграции в производственные системы. Особое внимание уделено использованию доверенных технологий ИИ и снижению вероятности утечки данных, что является критически важным аспектом при работе с чувствительной информацией. Архитектура должна быть гибкой и масштабируемой, чтобы удовлетворять разнообразным требованиям пользователей и адаптироваться к новым вызовам и технологиям. Общее представление функциональной архитектуры приведено на рисунке Б.2.

Разделение архитектуры на слои позволяет распределить функции с учетом различных аспектов работы системы синтеза, требований приватности и интеграции. Каждый слой охватывает ключевой функционал, делая акцент на определенных характеристиках системы.

Рекомендуется разделить систему на контуры, управляющие доступом к данным и отдельным модулям:

- внутренний контур: внутренняя часть системы, направленная на обработку исходных данных;

- партнерский контур: промежуточный уровень, где доступ строго контролируется, но доступ к реальным данным запрещен;

- публичный контур: условно-открытый доступ, ограниченный доступом к моделям и алгоритмам; используется только для конечных результатов.

Примечание — Разделение на контуры может уточняться в зависимости от требований к доступу при реализации системы синтеза.

Полное представление эталонной функциональной архитектуры представлено на рисунке Б.2.

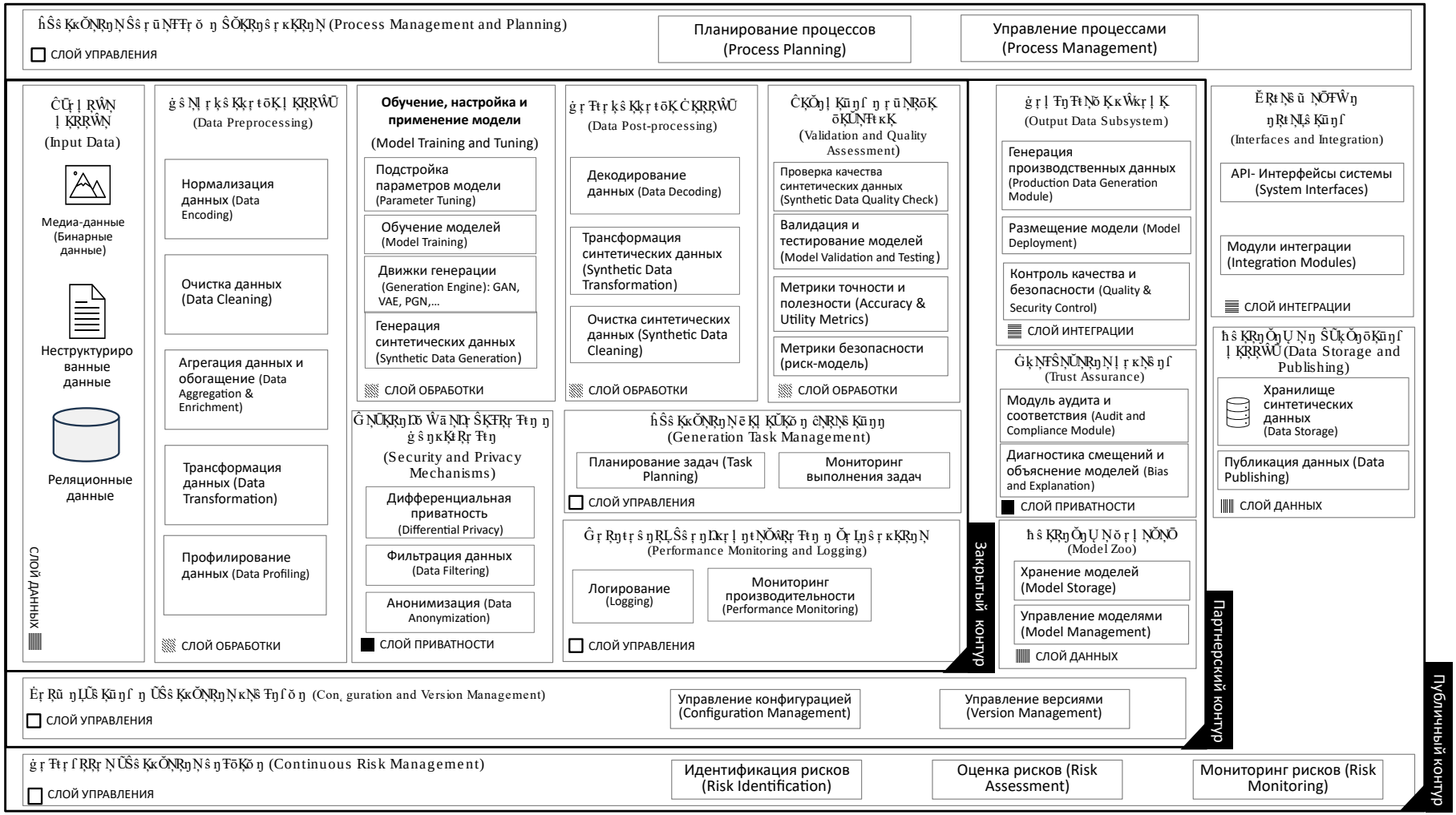


Рисунок Б.2 — Функциональная архитектура синтеза данных

Основные рекомендации для слоев функциональной архитектуры и контуров системы представлены в Б.2.1 (см. Таблицу Б.1).

Б.2.1 Слой данных

Слой данных отвечает за надежное и эффективное хранение всех данных, связанных с процессом синтеза, включая исходные, синтетические данные, а также модели, используемые для их генерации. Слой также управляет метаинформацией и гарантирует соблюдение всех требований к приватности и уровню доверия (см. таблицу Б.1).

Таблица Б.1 – Подсистемы слоя данных

Подсистема	Описание	Характеристики	Рекомендации по проектированию
Входные данные (Input Data)	Отвечает за сбор и хранение исходных данных	Поддержка различных типов данных (структурированные, полуструктурированные, неструктурированные данные). Процессы: DAT1, INF1, PLN1	Рекомендуется разработка интерфейсов для интеграции с различными источниками данных
Хранилище синтетических данных	Обеспечивает хранение и публикацию синтетических данных	Поддержка уровня доверия синтеза данных, разграничение доступа. Процессы: GEN1, PUB1, FBK1, KNW1, ETH1, DIS1	Учитывать требования приватности, в том числе контроль доступа
Хранилище моделей (Model Zoo)	Управление ЖЦ моделей генерации данных	Версионный контроль, обновление моделей. Процессы: MOD1, TST1, DEP1, DIS2	Внедрение системы управления конфигурацией для эффективного контроля версий

Слой данных рекомендуется проектировать с учетом обеспечения высокого уровня доверия к использованным технологиям синтеза. Особое внимание следует уделить разработке механизма управления доступом, учитывая различные контуры доступа (закрытый, публичный, партнерский). Рекомендуется использование распределенных систем хранения для повышения отказоустойчивости и масштабируемости.

Рекомендации по организации доступа:

- входные данные находятся в закрытом контуре с максимальным уровнем ограничения доступа; слой данных частично соответствует инфраструктурному уровню в эталонной архитектуре больших данных, описанному в ГОСТ Р ИСО/МЭК 20547-3;

- хранилище синтетических данных может находиться в публичном контуре, что требует реализации механизмов авторизации в зависимости от дизайна архитектуры приватности.

Б.2.2 Слой обработки данных

Слой обработки данных отвечает за полный цикл подготовки, трансформации и синтеза данных, включая последующую валидацию и оценку их качества. Слой обработки является ядром системы синтеза данных, предоставляя все необходимые функции для создания качественных и полных синтетических данных, список подсистем представлен в таблице Б.2.

Таблица Б.2 – Подсистемы слоя обработки данных

Подсистема	Описание	Характеристики	Рекомендации по проектированию
Предобработка данных (Data Preprocessing)	Подготовка данных для синтеза, включая нормализацию и очистку данных	Поддержка различных типов данных с обеспечением работы в закрытом контуре. Процессы: DAT1, REQ1, MOD1	Обеспечение гибкости для работы с различными форматами данных
Обучение и настройка модели (Model Training and Tuning)	Разработка, настройка и обучение моделей синтеза	Обработка исходных данных, настройка гиперпараметров. Процессы: MOD1, TST1, GEN1	Использование автоматизированных средств для подстройки параметров моделей
Постобработка данных (Data Post-processing)	Обработка синтетических данных после их синтеза	Включает декодирование, трансформацию и проверку данных на соответствие требованиям к уровню доверия к технологиям синтеза. Процессы: MOD1, PPS1, TST1, VAL1, PRI1	Внедрение механизмов очистки данных и устранения аномалий
Валидация и оценка качества (Validation and Quality Assessment)	Оценка качества и уровень приватности синтетических данных	Оценка метрик точности, полезности и вероятности утечки исходной информации. Процессы: VAL1, DBG1, TST1, GEN1, PPS1	Внедрение системы автоматизированного тестирования и оценки качества моделей

Общие рекомендации по реализации слоя обработки данных:

- при реализации слоя обработки данных следует уделить особое внимание обеспечению приватности данных на всех стадиях их обработки;

- рекомендуется использование современных методов МО и ИИ для повышения качества синтетических данных, внедрение механизмов автоматизированного контроля качества данных с учетом требований по уровню доверия методов синтеза, а также регулярного обновления моделей синтеза.

Б.2.3 Слой интеграции

Слой интеграции отвечает за взаимодействие системы синтеза данных с внешними системами и пользователями, обеспечивая управляемый доступ к синтетическим данным и моделям, а также управление публикацией данных и контролем доступа. В таблице Б.3 представлены основные подсистемы слоя интеграции.

Таблица Б.3 – Подсистемы слоя интеграции

Подсистема	Описание	Характеристики	Рекомендации по проектированию
Вывод данных (Output Data Subsystem)	Финальная обработка и генерация данных для производственного использования	Обработка и предоставление синтетических данных для использования с соблюдением требований качества и уровня доверия к использованным технологиям синтеза. Процессы: GEN1, DEP1, VAL1, QAS1	Учет требований реальных бизнес-процессов при проектировании систем публикации
Интерфейсы и интеграция (Interfaces and Integration)	Взаимодействие с внешними системами через модули интеграции	Поддержка стандартов обмена и модульная интеграция для обеспечения гибкости и предотвращения утечки данных при взаимодействии с внешними системами. Процессы: DEP1, PUB1, FBK1, OPR1	Следует убедиться в совместимости интерфейсов с существующими системами

Общие замечания по реализации слоя интеграции:

- слой интеграции следует проектировать с учетом требований совместимости с различными внешними системами и платформами, уделяя особое внимание обеспечению приватности и генерации полностью синтетических данных;

- особое внимание следует уделить стандартам обмена информацией и обеспечению качества полностью синтетических данных, передаваемых через открытые интерфейсы. Необходимо также предусмотреть механизмы масштабируемости и адаптации к изменяющимся бизнес-потребностям.

Примечания

1 Интерфейсы и модули интеграции обеспечивают защищенное взаимодействие между компонентами системы через внешние интерфейсы и интеграционные модули, сохраняя требования к уровню доверия к использованным технологиям синтеза.

2 Финальная обработка включает адаптацию синтетических данных для передачи конечному потребителю: преобразование в требуемый формат (CSV, JSON, Parquet и др.), формирование сопроводительных метаданных, упаковку и подготовку к публикации. В отличие от постобработки данных в закрытом контуре, которая направлена на техническую подготовку данных после синтеза (декодирование, трансформацию, устранение аномалий), финальная обработка ориентирована на требования внешних систем и потребителей и выполняется над уже проверенными данными без доступа к исходной информации.

Б.2.4 Слой приватности

Слой приватности в системе синтеза данных играет ключевую роль в реализации доверенных технологий ИИ на всех стадиях, включая обеспечение приватности и предотвращение несанкционированного доступа. Слой также поддерживает функции, направленные на уровень доверия к технологиям синтеза и предотвращение утечек данных, а также минимизацию восстановление исходных данных. Особое внимание уделено механизмам RDP и анонимизации, которые существенно снижают вероятность восстановления исходных данных и раскрытия ПДн в процессе синтеза. Слой приватности также должен обеспечивать доверие пользователей к системе через механизмы аудита, прозрачности и объяснимости моделей (см. таблицу Б.4).

Таблица Б.4 – Подсистемы слоя приватности

Подсистема	Описание	Характеристики	Рекомендации по проектированию
Механизмы приватности (Privacy Mechanisms)	Обеспечивает уровень доверия к технологиям синтеза на всех стадиях синтеза и минимизирует вероятность утечек данных	Включает RDP, фильтрацию и анонимизацию. Процессы: RMS1, TRST1, PRI1	Следует реализовать механизмы добавления контролируемого шума, фильтрации и маскировки данных для снижения вероятности утечки данных
Обеспечение доверия (Trust Assurance)	Поддерживает объяснимость решений моделей и отсутствие предвзятости в синтетических данных	Включает модули аудита и диагностики смещений. Процессы: TRST1, IMR1, QAS1, ETH1	Следует использовать методы LIME и SHAP для объяснения решений моделей и выявления смещений

Общие рекомендации по реализации слоя приватности:

- при проектировании слоя приватности необходимо предусмотреть изоляцию чувствительных данных внутри защищенного контура с ограниченным доступом;

- все механизмы снижения вероятности утечки данных должны соответствовать современным стандартам надежной обработки данных, учитывая требования ГОСТ Р 59347. Также следует внедрить системы мониторинга обработки данных, которые обеспечат оперативное выявление и устранение проблем.

Б.2.5 Слой управления системой

Слой управления системой в архитектуре синтеза данных включает подсистемы, ответственные за координацию и управление процессами синтеза данных, поддержание их бесперебойной работы и мониторинг всех стадий. Слой управления охватывает планирование и управление задачами генерации данных, мониторинг производительности системы, ведение журналов событий, управление конфигура-

цией и версиями компонентов, а также управление вероятностями утечек данных. Основной целью реализации слоя управления является обеспечение стабильной и предсказуемой работы системы, что особенно важно для выполнения критически важных задач, связанных с синтезом данных (см. таблицу Б.5).

Таблица Б.5 – Подсистемы слоя управления системой

Подсистема	Описание	Характеристики	Рекомендации по проектированию
Управление процессом синтеза и планирование (Process Management and Planning)	Координация процессов синтеза данных на всех стадиях	Планирование ресурсов и управление сроками. Процессы: PMG1, PLN1, CHG1, INF1, HRM1	Интеграция с системами планирования ресурсов для оптимального управления процессами
Управление задачами генерации (Generation Task Management)	Управление задачами синтеза данных, их очередностью и статусом	Мониторинг выполнения задач, корректировка планов. Процессы: PLN1, CHG1, INF1, CFG1, DOC1, DOC2	Внедрение автоматизированных систем управления задачами для повышения эффективности работы системы
Мониторинг производительности и логирование (Performance Monitoring and Logging)	Мониторинг ключевых показателей производительности и ведение журналов событий	Сквозная инфраструктура для отслеживания и анализа работы системы. Процессы: INF1, CFG1, OPR1	Внедрение механизмов мониторинга в реальном времени для отслеживания производительности системы
Конфигурация и управление версиями (Configuration and Version Management)	Управление конфигурацией и версиями всех компонентов системы	Обновление, контроль версий. Процессы: INF1, CFG1, OPR1	Использование централизованных систем управления версиями для предотвращения ошибок и повышения стабильности системы
Постоянное управление вероятностями утечек данных (Continuous Privacy Management)	Идентификация, оценка и постоянный мониторинг вероятностей утечки данных на всех стадиях синтеза	Оценка вероятности и воздействия факторов утечек данных. Процессы: RMS1, ETH1, TRST1, AUD1, PRI1, FBK1	Внедрение регулярного мониторинга утечек данных и механизмов их смягчения и предотвращения

Общие замечания по реализации слоя управления системой:

- система управления должна обеспечивать интеграцию всех процессов синтеза данных с централизованной системой мониторинга и управления вероятностями утечек данных;

- все компоненты управления должны быть тесно интегрированы с системой ведения журналов, что обеспечит оперативное реагирование на сбои и отклонения от нормальной работы;

- необходимо уделить особое внимание управлению конфигурациями и версиями компонентов системы для предотвращения ошибок и повышения стабильности системы.

Б.2.6 Указания к использованию функциональной архитектуры

Функциональная архитектура системы синтеза данных на основе ИИ предназначена для предоставления структурированного подхода к разработке, внедрению и управлению системой синтеза данных. Для эффективного использования архитектуры рекомендуется:

- идентификация требований: определить требования к системе на основе бизнес-целей и потребностей пользователей; учитывать требования к уровню доверия к методам синтеза, приватности, масштабируемости и производительности;

- декомпозиция системы: разделить проектируемую систему на подсистемы и модули в соответствии с предложенной архитектурой; убедиться, что каждый модуль имеет четко определенные функции и описанные взаимодействия с другими модулями;

- обеспечение приватности: внедрить механизмы предотвращения утечки данных на всех уровнях системы; гарантировать, что вероятность утечки данных минимальна при обработке и хранении данных;

- мониторинг и управление: настроить процессы мониторинга и управления для обеспечения бесперебойной работы системы и ее соответствия установленным требованиям;

- адаптация и масштабируемость: спроектировать систему с возможностью адаптации и масштабирования для будущих изменений и расширений;

- документирование и обучение: поддерживать актуальность документации системы и проводить обучение пользователей и администраторов правильному использованию и управлению системой; обеспечить доступность документации для всех заинтересованных сторон.

Примечание – Практическое применение архитектуры может включать сценарии использования больших данных в соответствии с ГОСТ Р 59926. Синтетические данные могут использоваться в деятельности государственных органов и коммерческих организаций для форматного контроля информационных систем, поддержки систем МО, а также для обработки и анализа данных.

Б.3 Проектирование архитектуры приватности

Системы синтеза данных представляют собой сложные распределенные информационные системы, которые обрабатывают и хранят большие объемы данных. Системы синтеза включают множество компонентов и стадий, которые подвержены различным проблемам утечки данных (см. [4]). Основные цели внедрения механизмов предотвращения утечки данных включают обеспечение минимально приемлемого уровня контроля данных на всех стадиях их обработки в соответствии с требованиями нормативных документов.

При проектировании архитектуры приватности систем синтеза необходимо учитывать следующие ключевые особенности (см. [4]):

- обработка и хранение больших объемов данных;
- интеграция различных компонентов и сервисов;
- многоступенчатая обработка данных: от получения исходных данных до генерации синтетических наборов.

Необходимо обратить внимание на то, что системы синтеза данных часто работают с чувствительной информацией. Несмотря на то, что синтетические данные создаются искусственно, они могут сохранять следы исходной информации, что влечет за собой вероятность утечки данных.

Ключевые проблемы синтеза данных, связанные с утечкой данных:

- присутствие квазиидентификаторов, которые могут быть использованы для идентификации личности при наличии дополнительной информации;
- наличие выбросов, которые могут выделяться и потенциально раскрывать исходную информацию.

Учитывая эти проблемы, проектирование архитектуры приватности должно быть приоритетным, с особым акцентом на предотвращение утечек данных через использование методов анонимизации, RDP и других современных доверенных технологий ИИ.

Б.3.1 Злоумышленники и атаки на систему синтеза данных

Для эффективной схемы синтеза данных с точки зрения снижения вероятности утечки данных необходимо учитывать типы злоумышленников и виды атак, которым может подвергаться система.

Типы доступа к данным злоумышленников:

- «белый ящик»: злоумышленник с полным доступом к системе, часто это инсайдер, например сотрудник компании с привилегиями для работы с данными; злоумышленник такого типа может анализировать внутренние механизмы системы и использовать свои привилегии для извлечения или изменения данных, что значительно увеличивает вероятность утечек данных и компрометации;
- «серый ящик»: злоумышленник с ограниченным доступом к системе, как правило, партнер или подрядчик, имеющий частичную информацию о внутренней структуре; такой злоумышленник может использовать свои знания для проведения целевых атак, например атак на идентификацию данных;
- «черный ящик»: злоумышленник, не имеющий доступа к системе, использует только внешние наблюдения и публичные данные; в этой роли могут выступать внешние агрессоры, стремящиеся нарушить работу системы через доступные интерфейсы и публичные данные.

К основным типам атак относятся:

- утечка данных: злоумышленник получает доступ непосредственно к исходной информации на любой стадии процесса синтеза данных;

Пример — Доступ к промежуточным данным, содержащим чувствительную информацию;

- атака идентификации: злоумышленник пытается выделить конкретную запись в синтетическом наборе данных, что может привести к раскрытию личности и нарушению приватности;

- атака на выводы: использование злоумышленником доступной информации для вывода скрытых данных или моделей, что может привести к раскрытию исходных данных, использованных для обучения модели;

- атака связывания: попытка злоумышленника сопоставить синтетические данные с внешними наборами данных для выявления совпадений; если синтетические данные недостаточно защищены, это может привести к идентификации личности;

- отравление данных: введение вредоносных данных в процесс генерации синтетических данных, что приводит к созданию некорректных или предвзятых данных, искажая результаты анализа;

- атака на принадлежность: определение злоумышленником того, является ли конкретная запись частью исходного набора данных, что может раскрыть присутствие или отсутствие конкретных данных в исходном наборе.

- атака на инверсию модели: восстановление злоумышленником исходных данных или их характеристик через доступ к модели или ее выводам; успешная атака может привести к утечке данных, использованных для обучения модели.

Для обеспечения необходимого уровня доверия системы синтеза данных следует использовать методы обеспечения приватности, такие как RDP, анонимизация, фильтрация данных и мониторинг на предмет возможных атак.

Б.3.2 Практическая реализация доверенных технологий

Практическая реализация доверенных технологий ИИ в системах синтеза данных играет ключевую роль в обеспечении уровня доверия к технологиям синтеза. Помимо стандартных методов, таких как контроль доступа, шифрование, аудит и мониторинг, особое внимание необходимо уделить методам, адаптированным для работы с синтетическими данными. В таблице Б.6 приведены дополнительные методы и технологии, примеры их реализации, а также их особенности и ограничения, включая учет переобученности модели и влияния объема данных.

Таблица Б.6 – Доверенные технологии ИИ и основанные на них методы синтеза

Метод	Описание	Реализация	Особенности и ограничения
Анонимизация	Удаление или маскирование идентифицирующей информации для предотвращения восстановления личности на основе данных	Применение k-анонимности для предотвращения утечки данных медицинских записей; использование l-разнообразия и t-близости для увеличения разнообразия в чувствительных атрибутах	Анонимизация может снижать точность данных и их полезность. Методы могут быть недостаточны при наличии квазиидентификаторов или выбросов. Эффективность уменьшается

Окончание таблицы Б.6

Метод	Описание	Реализация	Особенности и ограничения
			при большом объеме данных с высокой корреляцией
RDP	Добавление шума к данным для минимизации влияния отдельных записей на результаты обработки, соблюдая бюджет приватности (см. В.2)	Применение RDP в глубоком обучении с контролем бюджета приватности особым модулем (см. В.2)	Необходим тщательный баланс между уровнем шума и полезностью данных. Контроль бюджета приватности важен для сохранения ценности данных
Фильтрация данных	Удаление или маскирование данных, которые могут привести к утечке данных (см. 7.3.2)	Фильтрация выбросов в финансовых данных до их синтеза; маскирование идентифицирующих признаков в текстовых данных перед публикацией	Фильтрация может снижать качество данных и усложнить их обработку. Также возможно снижение эффективности при больших объемах данных, когда идентификация возможна через косвенные признаки
Оценка вероятности утечек данных	Оценка вероятностей утечек данных и компрометации системы на основе количественных метрик (см. ПНСТ 1066—2026)	Использование вероятностных моделей на основании оценок выбранного бюджета приватности или симуляции атак	Требуется регулярное обновление метрик и адаптация к новым проблемам приватности. Недооценка таких проблем может привести к утечке данных
Контроль переобучения модели	Минимизация вероятности переобучения модели для предотвращения утечек данных, которые могут быть восстановлены через синтетические данные (см. В.2.5)	Применение регуляризации и методов проверки на переобучение при обучении модели синтеза данных	При переобучении модель может запоминать исходные данные, что увеличивает вероятность утечек данных. Необходимо следить за балансом между качеством синтетических данных и переобучением
Управление объемом данных	Оценка влияния объема данных на эффективность методов снижения вероятности утечки данных, особенно RDP (см. В.2.5)	Регулирование объема данных для уменьшения вероятности истощения бюджета приватности, например ограничение размера синтетических выборок при создании данных с RDP	Чем больше данных обрабатывается, тем выше вероятность утечки данных и истощения бюджета приватности, что может снижать уровень доверия к технологиям синтеза

Рекомендации по реализации методов обеспечения уровня доверия синтеза:

- интеграция методов: методы обеспечения уровня доверия необходимо интегрировать в общую архитектуру системы синтеза, чтобы обеспечить многоуровневый подход к надежности создания полных синтетических данных на всех стадиях синтеза;

- адаптация к контексту: методы должны адаптироваться в зависимости от типа данных и уровня вероятности утечки данных; например, в системах с высокочувствительными данными акцент следует делать на RDP и управлении объемом данных;

- учет переобучения: при обучении моделей синтеза данных необходимо внедрить контроль переобучения, чтобы предотвратить запоминание исходных данных;

- регулирование объема данных: для систем с большим объемом данных следует предусмотреть эффективные механизмы управления объемом информации, чтобы не истощать бюджет приватности и сохранять стандарты обработки данных;

- регулярное обновление: методы обеспечения уровня доверия должны регулярно пересматриваться и обновляться с учетом новых проблем приватности и законодательных требований, что обеспечит актуальность и надежность системы.

Примечание — Обновление методов обеспечения уровня доверия и контроль за их эффективностью позволят снизить вероятность утечек данных и повысить доверие пользователей к системе синтеза данных.

Б.3.3 Типовая архитектура приватности

Архитектура приватности в системах синтеза является ключевым компонентом при разработке ИИ для генерации синтетических данных. Основная цель этой архитектуры — обеспечить отсутствие утечки данных во время обработки данных. Она основывается на стандартах и лучших практиках в области системной инженерии и безопасной обработки информации, таких как ГОСТ Р 59347, ГОСТ Р 59330 и ГОСТ Р 57100.

В типовой архитектуре обработки данных рекомендуется выделять три следующих основных контура, соответствующих различным уровням приватности данных и потенциальным типам злоумышленников:

- внутренний контур: отвечает за обработку наиболее чувствительных данных (непосредственно исходных данных), такой контур защищает от инсайдеров и предполагает максимальные меры по ограничению доступа к данным;

- партнерский контур: обеспечивает партнерам доступ к необходимым данным с минимальными вероятностями утечки данных; потенциальные злоумышленники с ограниченным доступом, такие как «шпионы», имеют доступ только к частичной и ограниченной информации;

- публичный контур: защищает данные и интерфейсы, доступные для внешних пользователей, обеспечивая сохранение свойств полностью синтетических данных при взаимодействии с внешними системами и предотвращая доступ к данным со стороны внешних злоумышленников («агрессоров»).

Архитектура приватности также предусматривает реализацию процессов непрерывного аудита, мониторинга и управления вероятностями утечки данных, включающих следующие меры (см. рисунок Б.3):

- регулярные проверки и обновление механизмов поддержки доверенных технологий ИИ для поддержания актуальности системы в условиях изменяющегося технологического ландшафта;

- мониторинг событий безопасности в реальном времени для быстрого выявления и устранения потенциальных утечек данных или инцидентов;

- оценка и управление проблемами утечки данных с учетом модели процессов, включая регулярные пересмотры методов оценки вероятности утечек данных и оценку возможных проблем.

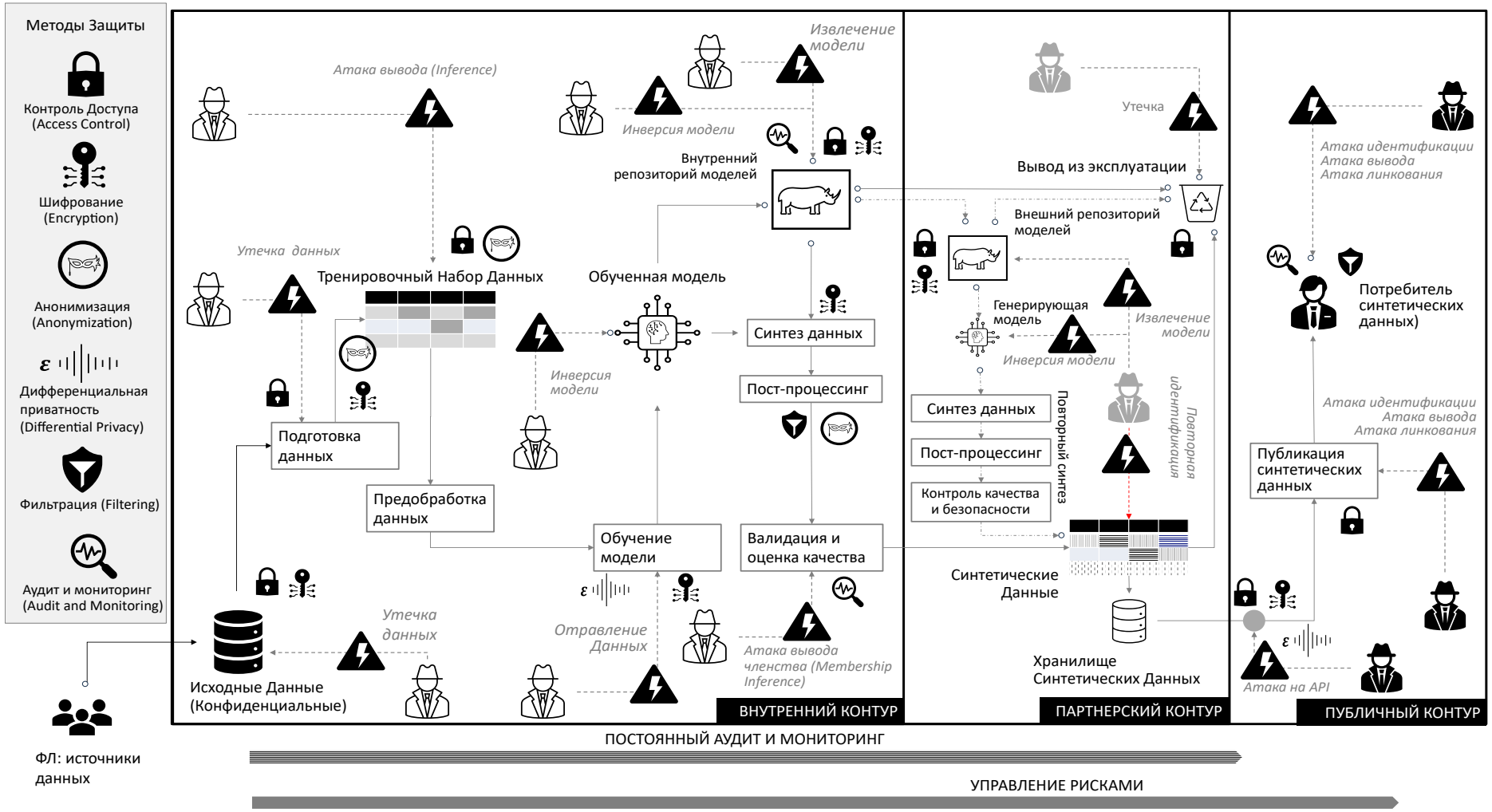


Рисунок Б.3 — Архитектура приватности системы синтеза

В таблице Б.7 представлены основные факторы, учтенные в данном представлении.

Таблица Б.7 – Компоненты архитектуры приватности системы синтеза

Компонент	Проблема	Методы
1 Исходные данные	Утечка данных	Контроль доступа, шифрование
2 Подготовка данных	Утечка данных	Контроль доступа, шифрование, анонимизация
3 Тренировочный набор данных	Атака на выводы	Шифрование, анонимизация
4 Обучение модели	Отравление данных	RDP, шифрование
5 Размещение обученной модели	Инверсия модели	Контроль доступа, шифрование
6 Внутренний репозиторий	Инверсия модели, извлечение модели	Аудит, контроль доступа, шифрование
7 Синтез данных и постпроцессинг	Утечка информации, атака линкования	Фильтрация, анонимизация, шифрование, контроль доступа
8 Внешний репозиторий	Извлечение модели, инверсия модели	Контроль доступа, шифрование
9 Валидация и оценка качества	Атака на принадлежность, утечка данных	Шифрование, RDP
10 Хранилище синтетических данных	Утечка данных, атака идентификации, атака на выводы, связывающая атака	Контроль доступа, шифрование
11 Публикация, интерфейсы	Атака на интерфейсы, человек посередине (MITM)	Шифрование, доступ к интерфейсам (HTTPS, OAuth 2.0, JWT)

Б.3.4 Применение архитектуры приватности

Архитектура приватности должна быть интегрирована на всех стадиях ЖЦ системы синтеза данных, включая как технические, так и организационные меры, обеспечивающие уровень доверия к технологиям синтеза и предотвращение утечек данных. Реализация данной архитектуры требует комплексного подхода, включающего следующие ключевые аспекты:

- проектирование: на стадии проектирования системы синтеза необходимо заложить следование регламентированным процессам, включающим меры по стандартной обработке исходных данных и контролю доступа; это обеспечивает уровень доверия к технологиям синтеза на всех уровнях системы, начиная с самого начала разработки. Также необходимо предусмотреть механизмы контроля объема данных, чтобы он соответствовал установленному бюджету приватности;

- разработка: в процессе разработки системы и моделей необходимо применять методы с гарантиями приватности, такие как RDP, а также механизмы предотвращения переобученности моделей; тестирование должно включать оценку вероятности

утечки данных и соответствие объема данных бюджету приватности для обеспечения надежной системы синтеза полностью синтетических данных;

- эксплуатация: во время эксплуатации систем синтеза данных требуется непрерывный мониторинг и управление приватностью; это включает оценку и управление вероятностями утечек данных, контроль переобученности моделей и ограничение числа повторных генераций данных на основе ранее обученных моделей; необходимо также соблюдать меры по ограничению доступа, особенно при работе с исходными данными, и вести постоянный аудит уровня вероятности утечки данных;

- вывод из эксплуатации: на стадии вывода системы синтеза данных из эксплуатации необходимо обеспечить безопасное удаление данных и деактивацию системы, минимизируя вероятности утечек данных и несанкционированного доступа; все данные должны быть удалены в соответствии с установленными протоколами, а доступ к системе должен быть полностью прекращен.

Б.4 Проектирование архитектуры интеграции

При проектировании архитектуры интеграции для систем синтеза данных необходимо учитывать несколько ключевых аспектов, которые обеспечат эффективность и безопасность интеграционных процессов. Выбор компонентов архитектуры должен основываться на специфических потребностях системы, типе данных, которые необходимо интегрировать, а также требованиях к уровню доверия к технологиям синтеза и совместимости. Практические рекомендации по проектированию архитектуры интеграции приведены в Б.4.1.

Б.4.1 Описание метаинформации

Для эффективного управления синтетическими наборами данных необходимо обеспечить структурирование метаинформации, которая будет сопровождать эти данные. Основные требования к метаинформации включают:

- универсальность: метаинформация должна быть пригодной для различных типов данных, включая текстовые, табличные и медиафайлы;

- стандартизированный формат: использование общепринятых стандартов для описания метаданных, таких как Дублинское ядро (Dublin Core, описание интернет-ресурсов), [5] (геопространственная информация), ДатаЦит (Data Cite, метаданные научных и академических исследований);

- актуальность и точность: метаинформация должна быть актуальной и точно отражать содержимое синтетических данных;

- читаемость и доступность: метаинформация должна быть легко читаемой и доступной для пользователей и систем.

Пример метаинформации для публикуемого синтетического набора в текстовом и табличном формате (см. таблицу Б.8).

Таблица Б.8 – Структура метаинформации (текстовый и табличный форматы)

Элемент	Описание
Идентификация	
Название набора данных	Краткое и информативное название набора данных
Идентификатор	Уникальный идентификатор набора данных (например, DOI)
Описание	
Аннотация	Краткое описание содержимого и назначения набора данных
Ключевые слова	Список ключевых слов для облегчения поиска и классификации
Происхождение	
Источник данных	Описание источников данных, использованных для создания синтетического набора
Методы синтеза	Подробное описание методов и алгоритмов, использованных для генерации синтетических данных
Структура данных	
Формат данных	Формат данных (например, JSON, CSV, XML)
Схема данных	Описание структуры данных, включая типы данных, названия столбцов и их описания
Размер данных	Объем данных, количество записей и размер файла
Дата/время	
Дата создания (Creation Date)	Дата создания набора данных
Период покрытия (Temporal Coverage)	Временной период, охватываемый данными
Правовые и общие аспекты	
Лицензия (License)	Условия использования и лицензия на набор данных
Смещенность и объяснимость (Bias and Explainability)	Описание любых вопросов, связанных со смещенностью, объяснимостью и использованием данных
Дополнительно	
Контактная информация (Contact Information)	Контактные данные авторов или создателей набора данных
Версионность (Versioning)	Информация о версиях набора данных и история изменений
Связанные публикации (Related Publications)	Список публикаций, связанных с набором данных

При описании метаинформации, сопровождающей синтетический набор для медиаформатов, рекомендуется придерживаться следующей структуры (см. таблицу Б.9).

Таблица Б.9 – Структура метаинформации в медиа форматах

Элемент	Описание
Идентификация	
Название медиафайла (Media Title)	Краткое и информативное название медиафайла
Идентификатор (Identifier)	Уникальный идентификатор медиафайла (например, DOI)
Описание	
Аннотация (Abstract)	Краткое описание содержимого и назначения медиафайла

Окончание таблицы Б.9

Элемент	Описание
Ключевые слова (Keywords)	Список ключевых слов для облегчения поиска и классификации
Происхождение	
Источник медиа (Source)	Описание источников медиа, использованных для создания синтетического файла
Методы синтеза (Synthesis Methods)	Подробное описание методов и алгоритмов, использованных для генерации синтетического медиафайла
Формат и структура	
Формат медиа (Media Format)	Формат медиафайла (например, JPEG, MP4, WAV)
Разрешение и качество (Resolution and Quality)	Технические характеристики, такие как разрешение изображения, частота кадров для видео или битрейт для аудио
Размер файла (File Size)	Объем файла
Дата/время	
Дата создания (Creation Date)	Дата создания медиафайла
Продолжительность (Duration)	Продолжительность видео- или аудиофайла
Правовые и общие аспекты	
Лицензия (License)	Условия использования и лицензия на медиафайл
Смещенность и объяснимость	Описание любых вопросов, связанных с использованием
Дополнительно	
Контактная информация (Contact Information)	Контактные данные авторов или создателей медиафайла
Версионность (Versioning)	Информация о версиях медиафайла и история изменений
Связанные публикации (Related Publications)	Список публикаций, связанных с медиафайлом

Б.4.2 Типовая архитектура интеграции

Рекомендуемая архитектура интеграции системы синтеза данных должна обеспечивать безопасное взаимодействие компонентов как внутри системы, так и с внешними системами, с обязательным учетом вероятности утечки данных и уровня доверия синтеза. На рисунке Б.4 приведена типовая схема архитектурного решения, демонстрирующая ключевые интеграционные механизмы в контексте ЖЦ системы синтеза данных.

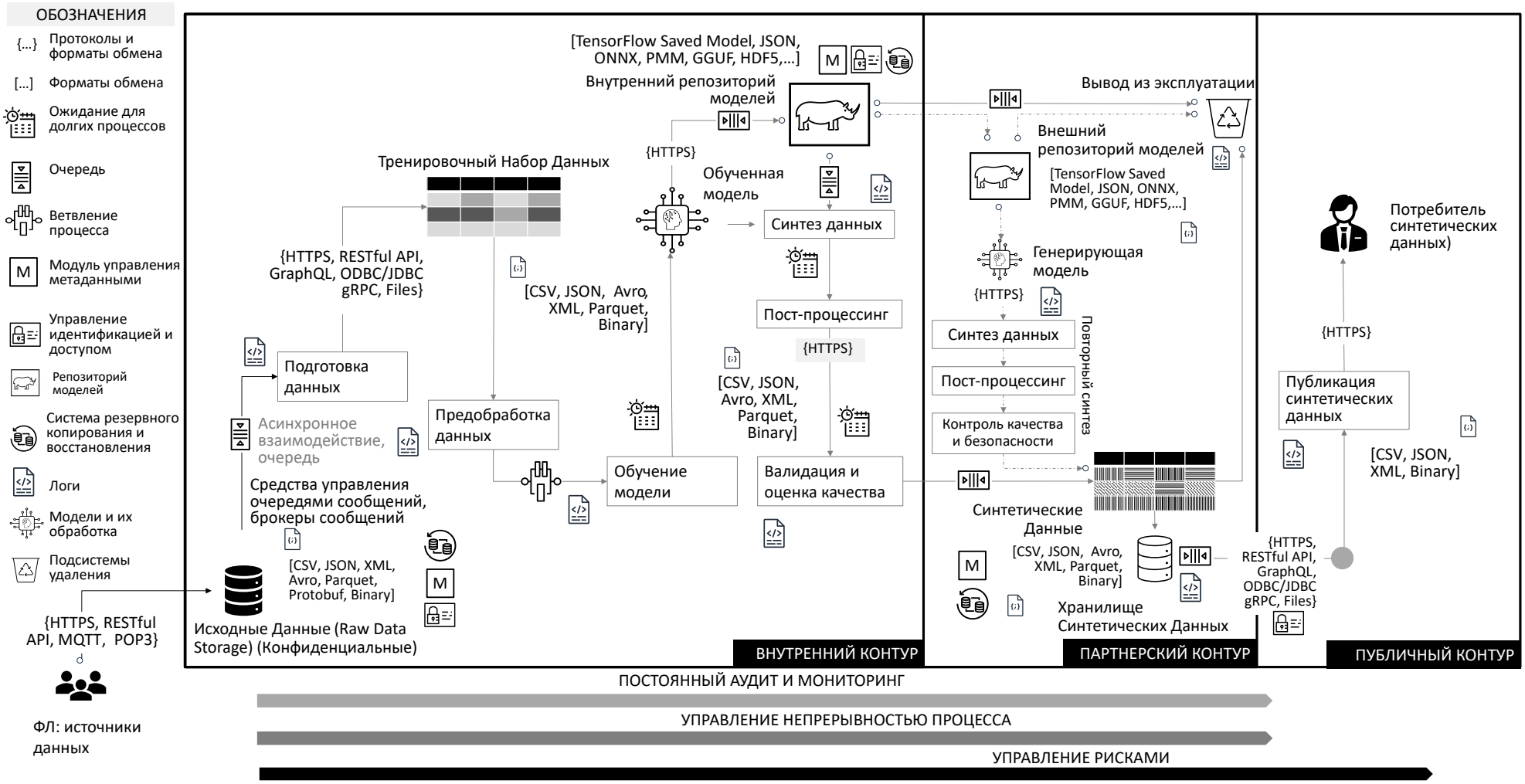


Рисунок Б.4 – Типовая архитектура интеграции системы синтеза

Рекомендации для эффективного использования интеграционной архитектуры:

- интеграция контуров безопасности: осуществляется через специально разработанные механизмы, обеспечивающие разделение потоков между внутренним, партнерским и публичным контурами с исключением утечек данных и несанкционированного доступа;

- асинхронные взаимодействия и управление очередями: использование асинхронных механизмов позволяет эффективно обрабатывать большие объемы данных при соблюдении требований приватности;

- разделение репозиториев моделей: внутренние и внешние репозитории разделяются для защиты конфиденциальных моделей и обеспечения контролируемого обмена с партнерами;

- управление метаинформацией: на каждой стадии синтеза должна сохраняться метаинформация о данных, их статусе и параметрах приватности для обеспечения прозрачности и доверия;

- подсистемы резервного копирования и логирования: необходимо предусмотреть модули для создания резервных копий и фиксации операций с целью восстановления и соблюдения требований приватности.

Б.4.3 Применение архитектуры интеграции

Архитектура интеграции должна быть реализована на всех стадиях синтеза данных:

- проектирование: выбор и интеграция мер обеспечения требуемого уровня приватности данных с учетом требований совместимости и масштабируемости (см. [4]);

- разработка: реализация и тестирование интеграционных модулей и интерфейсов для обеспечения надежности и уровня доверия к технологиям синтеза;

- эксплуатация: непрерывный мониторинг и управление интеграционными процессами, включая аудит и контроль за состоянием системы;

- вывод из эксплуатации: безопасное удаление данных и деактивация интеграционных модулей, с сохранением возможностей восстановления данных.

Приложение В
(справочное)
Технические особенности реализации синтеза данных

В.1 Общий подход к методам синтеза

Настоящее приложение содержит справочные сведения о технических аспектах реализации синтеза данных с использованием систем ИИ. Представлены основные подходы, применяемые для генерации синтетических данных на основе алгоритмов МО, включая методы, учитывающие требования к приватности, воспроизводимости и качеству выходных данных.

Методы синтеза данных являются ключевым элементом формирования качественного синтетического набора. Такие методы включают:

- алгоритмы и подходы к извлечению статистических моделей из исходных данных;
- стратегии построения необходимого уровня обобщения и обрывов связи с реальными объектами, обеспечивающие генерацию новых, реалистичных, но неперсонифицированных данных;
- механизмы контроля вариативности, стабильности и регуляризации модели генерации.

Выбор метода синтеза оказывает прямое влияние:

- на типы и структуру обрабатываемых данных (например, табличные, временные ряды, изображения);
- реалистичность, применимость и полезность создаваемых данных;
- уровень приватности, включая допустимость повторной генерации и устойчивость к атакам восстановления;
- вычислительные ресурсы, необходимые для обучения и генерации;
- надежность и устойчивость модели при изменении входных условий.

Описанные методы не являются исчерпывающими и могут быть дополнены другими подходами в зависимости от целей синтеза, ограничений в домене применения и предпочтительных механизмов контроля приватности. Основной процесс генерации синтетических данных остается структурно единым, но выбор метода может существенно влиять на промежуточные этапы, качество, и применимость результата.

В.2 Обзор основных алгоритмов синтеза данных

В.2.1 Генеративно-состязательные сети

GAN представляют собой класс алгоритмов МО, где система включает две нейронные сети: генератор и дискриминатор. Эти сети обучаются совместно в состязательном режиме, что позволяет каждой из них улучшать свои способности через взаимодействие. Принципиальная схема реализации генеративной сети представлена на рисунке В.1:

- генератор: задача генератора заключается в создании синтетических данных, которые максимально похожи на реальные (например, при использовании RDP генератор получает на вход случайный шум и преобразует его в синтетические данные,

которые пытается выдать за настоящие). Основная цель генератора — «обмануть» дискриминатор, заставив его принять синтетические данные за реальные;

- дискриминатор: задача дискриминатора заключается в различении реальных данных от синтетических. Он принимает на вход как реальные данные, так и данные, созданные генератором, и обучается их различать. Дискриминатор оценивает вероятность того, что данные являются настоящими, и стремится минимизировать ошибки в своих оценках.

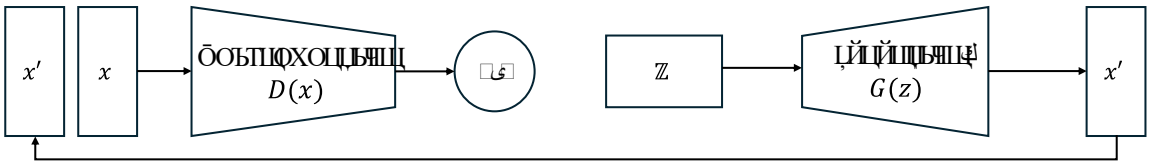


Рисунок В.1 — Принципиальная схема реализации GAN

Использование алгоритмов GAN для синтеза данных имеет следующие преимущества:

- высокая реалистичность генерируемых данных: GAN способны создавать данные, которые визуально или статистически очень похожи на реальные данные;

- возможность работы с различными типами данных: такая архитектура сети успешно применяется для генерации изображений, текста, аудио и других типов данных;

- адаптивное обучение: благодаря состязательной природе, генератор и дискриминатор постоянно улучшаются, что приводит к повышению качества синтезируемых данных.

В то же время, такая архитектура имеет ограничения:

- трудности в обучении: процесс обучения такой сети может быть нестабильным и сложным из-за необходимости поддержания баланса между генератором и дискриминатором;

- возможные нестабильности: GAN могут сталкиваться с проблемами, такими как затухание градиентов или захват, что приводит к ухудшению качества генерации;

- сложность настройки гиперпараметров: для достижения оптимальной производительности требуется тщательная настройка множества гиперпараметров.

В.2.2 Вариационные автокодировщики

Вариационные автокодировщики — класс алгоритмов МО, использующий вероятностный подход для моделирования сложных распределений данных. Архитектура вариационного автокодировщика состоит из двух ключевых компонентов: кодера и декодера, которые совместно обучаются для преобразования данных в латентное пространство и их восстановления обратно в исходное состояние. Принципиальная архитектура вариационного автокодировщика представлена на рисунке В.2.

Латентное пространство является основным понятием при применении архитектур вариационного автокодировщика и представляет собой область хранения данных в сжатом и упрощенном виде (абстрактной форме), которое моделируется вероятностно, что позволяет автокодировщику учиться генерировать данные, которые соответствуют распределению, наблюдаемому в исходных данных.

Основные компоненты вариационного автокодировщика:

- кодер: преобразует исходные данные во внутреннее представление в латентном пространстве в виде параметризованного распределения;

- декодер: восстанавливает данные из точек латентного пространства, стремясь к максимальному приближению к исходным данным.

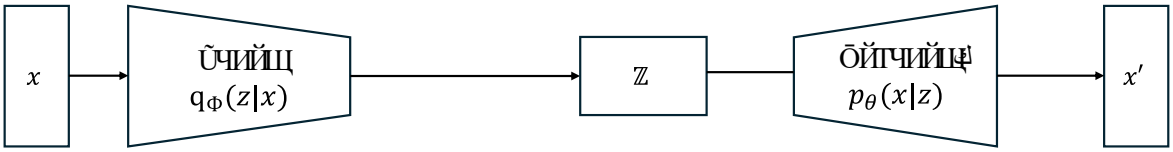


Рисунок В.2 — Архитектура вариационного автокодировщика

Использование архитектуры вариационного автокодировщика имеет следующие преимущества:

- стабильность: по сравнению с GAN, процесс обучения вариационный автокодировщик более устойчив, что упрощает его настройку и повышает предсказуемость результатов;

- генерация непрерывных данных: благодаря вероятностному подходу, вариационный автокодировщик может плавно интерполировать между различными точками в латентном пространстве, что позволяет генерировать непрерывные данные.

Синтез данных с помощью вариационных автокодировщиков имеет следующие ограничения:

- менее реалистичные данные: по сравнению с GAN, данные, сгенерированные вариационным автокодировщиком, могут выглядеть менее реалистично, поскольку вариационный автокодировщик стремится к максимизации правдоподобия, а не к максимальной правдоподобности для дискриминатора;

- сложность моделирования дискретных данных: вариационный автокодировщик лучше справляется с непрерывными данными и может испытывать трудности при работе с дискретными данными.

Примерами применения вариационного автокодировщика является синтез новых изображений на основе обучения на большом количестве существующих изображений, обнаружение аномалий в данных путем сравнения вероятности данных в латентном пространстве, восстановление поврежденных данных.

В.2.3 Вероятностные графические сети

Вероятностные графические сети — класс алгоритмов МО, предназначенный для генерации последовательных данных и временных рядов. Такие сети обучаются предсказывать следующие элементы последовательности на основе предыдущих, что позволяет им моделировать сложные временные зависимости в данных. Типовая организация вероятностной графической сети представлена на рисунке В.3.

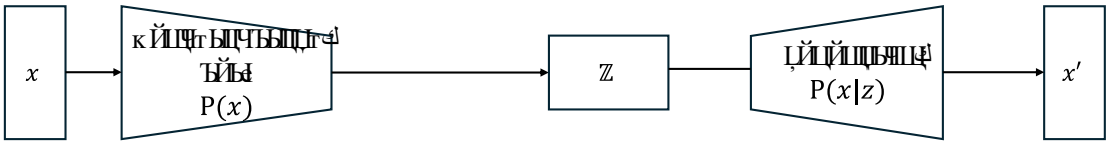


Рисунок В.3 — Архитектура синтеза с использованием вероятностной графической сети

Типичным представителем этого класса решений является байесовская сеть — вероятностная графическая модель, представляющая множество переменных и их условные зависимости с помощью направленного ациклического графа. Такой граф обладает набором полезных математических свойств, таких как топологическая сортировка, и удобен для моделирования зависимостей. В контексте вероятностных графических сетей байесовская сеть используется для моделирования зависимостей между последовательными элементами данных. Сеть состоит из следующих элементов:

- узлы: представляют переменные (например, элементы временного ряда);
- ребра: представляют условные зависимости между переменными.

Байесовская сеть строится путем определения структуры зависимостей между переменными и задания вероятностных распределений для этих зависимостей. Это позволяет байесовским сетям эффективно моделировать и предсказывать сложные временные зависимости в данных.

Преимущества использования предсказательных сетей:

- эффективность при работе с временными рядами: вероятные графические сети хорошо справляются с моделированием временных зависимостей, что делает их идеальными для работы с временными рядами и последовательными данными;
- улучшенная обработка последовательностей: способность предсказывать следующие элементы последовательности позволяет сети генерировать данные, которые сохраняют структуру и закономерности исходных данных.

Применение архитектуры вероятностной графической сети также имеет ограничения:

- такие архитектуры требуют большого объема данных для обучения;
- возможны ошибки предсказания на достаточно больших последовательностях данных: при работе с такими последовательностями точность предсказаний может снижаться, что приводит к накоплению ошибок;
- графовые сети подвержены специфическим атакам на граф, переобучению моделей, корректность синтеза на такой архитектуре должна тщательно проверяться и иметь гарантии приватности для обоснования применения.

Такого рода архитектуры часто применяются для генерации временных рядов и синтеза последовательности событий.

Примечание — В отличии от нейронных сетей, PGN обучаются без использования градиентных методов или методов обратного распространения ошибок, применение ограничено трудоемкостью их построения, а использование без строгих гарантий приватности, полученных, например, с помощью RDP не рекомендуется за пределами контура доступа к исходной информации.

В.2.4 Потоковые сети

Потоковые сети — это класс алгоритмов МО, предназначенный для моделирования вероятностных распределений данных с использованием обратимых преобразований.

В отличие от вероятностных генеративных сетей, которые фокусируются на последовательных данных и временных рядах, потоковые сети стремятся к точному моделированию распределений данных в их пространстве признаков. Основная идея такого подхода заключается в преобразовании сложного распределения данных в простое (как правило, нормальное) распределение через последовательность обратимых и дифференцируемых преобразований. Упрощенная архитектура потоковых сетей представлена на рисунке В.4.

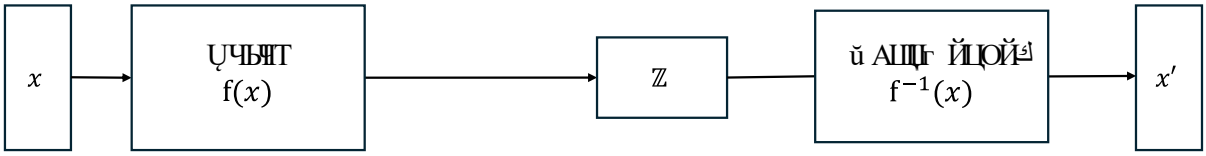


Рисунок В.4 — Архитектура потоковых сетей для синтеза данных

Преимущества потоковых сетей:

- точное моделирование распределений: потоковые сети способны моделировать сложные распределения данных с высокой степенью точности;

- возможность обратимости: каждое преобразование в потоковой сети является обратимым, что позволяет точно восстанавливать исходные данные из их латентного представления.

Ограничения потоковых сетей:

- высокие вычислительные затраты: потоковые сети требуют значительных вычислительных ресурсов для обучения и генерации данных из-за сложности обратимых преобразований;

- сложность в реализации: реализация потоковых сетей может быть технически затратна и плохо автоматизируется.

Примеры применения включают генерацию синтетических изображений с высокой степенью точности, генерацию видеоконтента, приложения виртуальной реальности.

В.2.5 Трансформеры

Трансформеры — это класс моделей МО, предназначенный для обработки последовательностей данных с использованием механизма самовнимания. Архитектура трансформеров позволяет учитывать зависимость каждого элемента последовательности от всех других элементов, что значительно повышает качество обработки данных по сравнению с традиционными рекуррентными нейронными сетями и сетями с долгой/краткосрочной памятью.

Трансформер состоит из двух основных компонентов: энкодера и декодера. В стандартной полной архитектуре трансформера используются оба компонента, однако существуют вариации, использующие только энкодер или только декодер:

- энкодер: включает слои самовнимания и полностью связанные слои, обеспечивая учет контекста каждого элемента входной последовательности;

- декодер: дополнительно использует внимание к выходам энкодера, позволяя формировать последовательности с учетом как предыдущих элементов, так и контекста.

Механизм самовнимания, позволяет модели вычислять взвешенную сумму всех элементов последовательности, где веса определяются важностью каждого элемента по отношению к другим, что позволяет трансформеру эффективно моделировать зависимости на больших расстояниях в последовательности данных. Упрощенно архитектура трансформера для синтеза данных представлена на рисунке В.5.

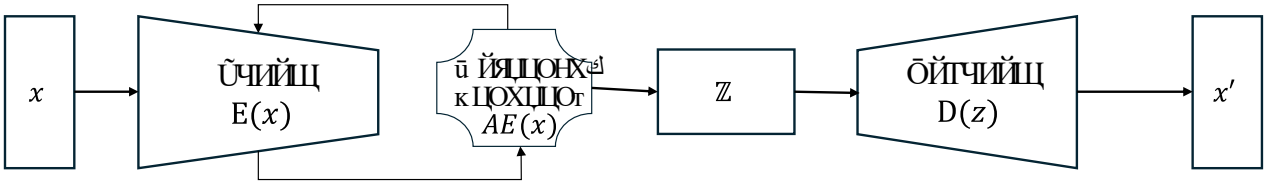


Рисунок В.5 — Архитектура трансформера в процессе синтеза с шумом

Архитектуру трансформеров применяют для реализации систем синтеза текста на основании ввода пользователя (prompt), так называемых больших языковых моделей. Такие модели обучены на больших объемах текстовых данных и синтезируют свободный текст, код, изображения (мультимодальные реализации).

Преимущества использования трансформеров:

- высокая эффективность обработки последовательностей: трансформеры могут эффективно обрабатывать длинные последовательности данных благодаря механизму самовнимания;

- масштабируемость: трансформеры могут быть легко масштабированы путем добавления дополнительных слоев;

- возможность параллельной обработки: в отличие от рекуррентных сетей и сетей с длинной/короткой памятью, трансформеры могут обрабатывать данные параллельно, что значительно ускоряет обучение и вывод.

Несмотря на расширение списка сценариев применения больших языковых моделей, трансформеры не лишены ограничений:

- высокие требования к вычислительным ресурсам: трансформеры требуют значительных вычислительных ресурсов для обучения и вывода, особенно в случае больших моделей;

- сложность настройки: настройка гиперпараметров и структуры модели может быть сложной задачей, требующей значительного опыта и знаний;

- ограничения в моделировании числовых зависимостей: трансформеры не предназначены в эталонной реализации к моделированию сложных корреляций и распределений на основе исходных данных, ограничиваясь преимущественно слабоструктурированными текстами;

- проблемы с уровнем доверия к технологиям синтеза на базе трансформеров: хотя предлагаются подходы с RDP и активно применяется фильтрация, приватность вывода трансформера не гарантирована.

В.3 Дифференциальная приватность

В.3.1 Гарантии приватности

Ранее среди методов снижения вероятности утечки данных была рассмотрена RDP. Техника RDP имеет значительное влияние (см., например, [6] на обеспечение приватности в процессе синтеза и надежной обработки данных, поэтому выделена дополнительно в настоящем стандарте.

RDP— это математический подход к снижению вероятности утечки данных, который позволяет проводить анализ и синтез данных, минимизируя вероятность утечки данных о конкретных индивидуальных записях (см. ПНСТ 1064—2026, пункт 3.3.5) (см. также [7]). Основная идея механизма заключается в добавлении контролируемого шума к данным или результатам вычислений, чтобы сделать невозможным определение присутствия или отсутствия любого отдельного индивида в наборе данных.

Примечание — Как формальная модель измерения приватности, модель дифференциальная приватность была определена в работе Дворк в виде условия для рандомизированного алгоритма M, принимающего набор данных D_i в качестве входных данных: алгоритм *M* считается **ϵ**-дифференциально приватным для всех наборов данных D_1 и D_2, которые отличаются одним элементом (т. е. данными для одного принципала данных), и для всех подмножеств S диапазона M, если:

$$\Pr[M(D_1) \in S] \leq e^{\epsilon} \cdot \Pr[M(D_2) \in S] + \delta, \qquad (\text{B}.1)$$

где ϵ контролирует уровень приватности (чем меньше ϵ, тем выше приватность), а δ — вероятность того, что механизм нарушает ϵ-дифференциальную приватность. *S* — это множество всех возможных выходов механизма *M* (пространство результатов).

Приведенное выше определение называют также (ϵ, δ)-дифференциальной приватностью или приблизительной дифференциальной приватности , иногда также рассматривая строгую ϵ-дифференциальную приватность (чистая дифференциальная приватность, Pure DP), в которой $\delta = 0$.

Среди других видов дифференциальной приватности распространение также имеет RDP — обобщение классической дифференциальной приватности (с использованием дивергенции Реньи для измерения различий между распределениями данных. В отличие от традиционной дифференциальной приватности (строгой дифференциальной приватности), которая использует параметр ε для контроля приватности, дивергенция Реньи использует параметр α (степень дивергенции Реньи) и вводит более гибкую метрику приватности. Базовое условие дифференциальной приватности, эквивалентное выражению (В.1) можно сформулировать как:

$$D_{\alpha}(\mathcal{M}(D_1)||\mathcal{M}(D_2)) \leq \epsilon(\alpha), \qquad (\text{B}.2)$$

где D_{α} – дивергенция Реньи, рассчитываемая по формуле

$$D_{\alpha}(P \parallel Q) = \frac{1}{\alpha - 1} \log\left(\sum_i P(i)^{\alpha} Q(i)^{1-\alpha}\right). \qquad (\text{B}.3)$$

Может быть отражено, что RDP эквивалентна (ϵ, δ)-дифференциальной приватности.

В рамках дифференциальной приватности приведенные выше математические определения называют также гарантиями конфиденциальности, под которыми пони-

маются формальные математические обещания, что участие или неучастие отдельного ФЛ в наборе данных практически не повлияет на выводы, сделанные из этого набора данных.

Большое значение для корректного применения дифференциальной приватности имеет выбранная единица приватности, определяющая объект, относительно которого имеется чувствительная информация: отдельная запись, группа записей, персона. Выбор единицы приватности определяет, как параметр ϵ будет интерпретироваться в контексте утечек данных. Чем больше данных охватывает единица приватности, тем выше должен быть уровень шума (меньше ϵ), так как каждая единица содержит больше информации.

В.3.2 Свойства дифференциальной приватности

Дифференциальная приватность обладает рядом полезных свойств, которые могут быть сформулированы как теоремы и доказаны строго:

- теорема композиции: утверждает, что несколько механизмов, удовлетворяющих дифференциальной приватности, которые применяются к одному и тому же набору данных, приводят к тому, что параметры приватности накапливаются линейно:

$$\epsilon_{total} = \sum_{i=1}^{m} \epsilon_i \,, \tag{В.4}$$

$$\delta_{total} = \sum_{i=1}^{m} \delta_i \,; \tag{В.5}$$

- теорема постобработки: утверждает, что любая операция, выполненная над данными, уже обработанными механизмом дифференциальной приватности, не ухудшает уровень приватности;

- свойство групповой приватности в дифференциальной приватности расширяет гарантии приватности с уровня одной записи на уровень группы записей: позволяет оценивать, как изменяются параметры приватности, когда необходимо обеспечить приватность не отдельных записей, а группы взаимосвязанных данных; свойство применяется, когда данные могут быть сгруппированы по k пользователям, транзакциям или другим логическим связям, в этом случае для ϵ-дифференциальной приватности можно записать условие:

$$\frac{\Pr[\mathcal{M}(D_1') \in S]}{\Pr[\mathcal{M}(D_2') \in S]} \leq e^{\epsilon k}; \tag{В.6}$$

- теорема о разделении источников формализует подход, при котором различные источники данных обрабатываются с помощью отдельных механизмов дифференциальной приватности: применяется для случая, когда данные поступают из нескольких независимых источников, и необходимо комбинировать результаты этих источников, сохраняя приватность каждой отдельной части данных; теорема показывает, что приватность каждого источника можно рассматривать независимо, без необходимости накладывать дополнительные ограничения на итоговый уровень приватности:

$$\epsilon_{total} = \max(\epsilon_1, \epsilon_2, \dots, \epsilon_n), \quad (\text{В.}7)$$

$$\delta_{total} = \max(\delta_1, \delta_2, \dots, \delta_n). \quad (\text{В.}8)$$

Примечание – Объем добавляемого шума является критическим параметром, поскольку его увеличение снижает качество данных и одновременно снижает вероятность утечки данных. Поэтому разработчики стремятся к экономному расходованию шума. Такую экономию позволяют обеспечивать дополнительные теоремы – продвинутой композиции и композиции гетерогенных механизмов, позволяющие более строго оценивать утечку данных по отношению к линейной аппроксимации.

В.3.3 Бюджет приватности

Бюджет приватности в контексте дифференциальной приватности представляет собой ограничение на количество информации, которое может быть раскрыто об отдельных данных через многократное использование алгоритмов с приватностью (см. [7]). В практическом смысле под бюджетом приватности понимают объем шума, добавляемого к данным, и отождествляют бюджет с параметром ϵ. Он определяет, насколько часто и насколько сильно можно взаимодействовать с набором данных, прежде чем утечка данных станет значительной. Ключевые характеристики управления бюджетом приватности:

- параметр ϵ определяет объем бюджета: чем он меньше, тем выше защита, но тем меньше допустимых вычислений;
- композиция означает, что каждый запуск алгоритма тратит часть бюджета, и при накоплении операций возрастает риск утечки данных;
- контроль утечки данных реализуется через ограничение числа операций, которые можно выполнить, не нарушая приватность индивидуальных записей;
- перераспределение бюджета позволяет гибко назначать значения ϵ в зависимости от приоритета вычислений;
- параметр δ, применяемый в некоторых случаях, ограничивает вероятность серьезной утечки данных и повышает гибкость управления.

Добавление шума, как правило, производится с использованием функций распределения Лапласа для строгой дифференциальной приватности или Гаусса – приблизительная дифференциальная приватность. Величина ϵ называется также бюджетом приватности, определяя, сколько «шума» добавляется к данным для достижения необходимой приватности. В процессе синтеза данных объем израсходованного шума контролируется специальным механизмом учета (accountant), реализуемым с помощью функций:

- метод PATE (Private Aggregation of Teacher Ensembles): использует ансамбль учителей, обученных на подмножествах данных, для агрегирования ответов с добавлением шума, обеспечивая контроль бюджета приватности;
- метод «Момент аккаунтов»: использует моменты распределений для более точного учета использования бюджета приватности в алгоритмах МО;
- механизм учета Реньи: основывается на концепции дивергенции Ренье для точного учета бюджета приватности.

В.3.4 Способы обеспечения дифференциальной приватности

В зависимости от способа управления приватностью различают два ведущих метода обеспечения дифференциальной приватности в процессе глубокого обучения (то есть при использовании глубоких нейронных сетей):

- ансамбль учителей PATE (Private Aggregation of Teacher Ensembles): использует обучение нескольких моделей на разных подмножествах и агрегирует их ответы с добавлением шума, обеспечивая высокую приватность при больших вычислительных затратах (см. [8]);

- метод стохастического градиента DP-SGD (Differentially Private Stochastic Gradient Descent): добавляет шум к градиентам на каждом шаге обучения, что снижает вычислительную нагрузку, но может уменьшать точность модели при высоком уровне приватности.

Примечание – Следует различать применение методов дифференциальной приватности в обучении нейронных сетей и общие подходы дифференциальной приватности. В частности, регулирование бюджета приватности через количество запросов при использовании механизмов дифференциальной приватности для запросов через интерфейсы отличается от его расхода при глубоком обучении. В случае нейронных сетей бюджет приватности расходуется по мере того, как модель обновляется в процессе обучения, что может быть сопоставлено вычислению градиентов как для DP-SGD. Связь между расходом бюджета и объемом исходных данных при синтезе данных рассматривается в В.3.5.

В.3.5 Влияние объема исходных данных

Размер исходного набора данных влияет на параметры дифференциальной приватности, такие как ε и δ. Адекватный объем данных необходим для достижения заданного уровня приватности и точности.

Объем исходного набора данных:

- минимальный объем данных: для обеспечения требуемого уровня приватности и точности минимальное количество записей *n* (минимальный размер исходного набора данных) должно удовлетворять условию (см. [1])

$$n \geq \frac{2}{\epsilon\delta} \cdot \sqrt{|F| \log\left(\frac{|F|}{\gamma}\right)}, \qquad \text{(В. 9)}$$

где ϵ — параметр приватности;
δ — допустимая погрешность;
γ — вероятность ошибки;
$|F|$ — количество доступных для набора данных линейных статистик.

- максимальный объем данных: для некоторых механизмов приватности, например RDP, оценка сверху для числа записей n дает:

$$n \leq \frac{\epsilon \cdot 2\sigma^2}{\alpha\Delta^2}, \qquad \text{(В. 10)}$$

где Δ — чувствительность данных по норме *L*2;

σ^2 — дисперсия шума;

α – число Реньи.

Для механизма Гаусса зависимость $\epsilon(n)$ задается формулой

$$\epsilon(n) = \frac{\epsilon_0}{\sqrt{n}}\ , \qquad (\text{В.}11)$$

где ϵ_0 – исходный бюджет приватности для эталонного набора из одной записи ($n = 1$), чем больше n, тем меньший шум требуется для обеспечения той же вероятности утечки данных;

- учет структуры данных: в данных с высокой корреляцией чувствительность Δ может увеличиваться, что потребует большего объема данных. Многократные запросы могут увеличивать утечку данных быстрее, что требует учета влияния каждого запроса на общий бюджет.

Примечания

1 Механизм Гаусса: зависимость (В.11) является упрощенной и может не точно отражать реальное поведение механизма при больших объемах данных.

2 Другие механизмы**:** существуют и другие механизмы дифференциальной приватности, для которых зависимости между ϵ и n могут иметь более сложный вид.

Если количество строк данных существенно меньше количества параметров модели, возникает вероятность переобучения. Для контроля этого процесса применяется специальная метрика — размерность Вапника-Червоненкиса (VC-размерность, VC_{dim}). Эмпирически VC-размерность можно оценить через соотношение:

$$VC_{dim} \approx L \cdot W\ , \qquad (\text{В.}12)$$

где L — количество слоев модели;

W — количество нейронов в слое нейронной сети.

Количество параметров модели N_P в этом случае должно удовлетворять условию:

$$N_P \leq VC_{dim} \cdot \log N\ , \qquad (\text{В.}13)$$

где N — количество строк исходного набора данных (или аналогичная характеристика).

Если это условие не соблюдается, модель может быть переобучена, что повышает вероятность утечек данных через атаки, такие как атака на членство.

Примечание – Приведенные в виде оценки для VC_{dim} являются приблизительными, некоторые методы ИИ, например, вероятностные графические сети (см. В.2.3), имеют ряд сложностей в выявлении переобучения и сложны для построения надежных методов синтеза.

В.3.6 Практическая реализация дифференциально приватности

Примеры реализации перспективных методов дифференциальной приватности при синтезе данных:

- PATE-GAN: используется метод PATE для обучения нескольких дискриминаторов (учителей) на различных подмножествах данных. Ответы учителей агрегируются с добавлением шума, чтобы обеспечить дифференциальную приватность;

- CTAB-GAN+: используется метод DP-SGD для модификации стандартного алгоритма стохастического градиентного спуска, добавляя шум к градиентам на каждом шаге обновления параметров модели;

- PrivBayes: применяет методы дифференциальной приватности для создания синтетических данных на основе байесовских сетей (вероятностные графические модели) с добавлением шума к параметрам модели.

В.4 Обогащение синтетических данных

Обогащение синтетических данных — это процесс добавления дополнительной информации к уже сгенерированным синтетическим данным. Этот процесс может быть необходим в случаях, когда синтетические данные недостаточно точны или детализированы, чтобы удовлетворить требованиям конкретных приложений. Обогащение данных позволяет повысить их качество, точность и полезность, обеспечивая более релевантные и ценные результаты для анализа и применения (см. [8]).

Несмотря на то, что синтез данных позволяет учесть нюансы исходных данных, могут возникать ситуации, требующие обновления уже сгенерированных данных:

- недостаток исходной информации может привести к чрезмерной обобщенности синтетических данных и потере важных деталей;

- специфические требования приложений требуют повышенной точности и детализации, особенно в медицине и финансах;

- обеспечение контекста достигается путем добавления дополнительной информации, повышающей практическую ценность данных.

В целях обогащения синтетических данных применяют несколько подходов, включая дополнительные ветвления в процесс синтеза данных, описанный в разделе 5.

В.4.1 Обогащение из внешних источников

Для добавления к синтетическим данным внешней информации применяют методы извлечения и обогащения в процессе генерации: генерация с аугментацией. На рисунке В.6 представлена принципиальная схема процесса синтеза с обогащением. Процесс генерации с аугментацией включает следующие этапы:

- извлечение информации: на этом этапе с использованием поисковых алгоритмов, методов регулярных преобразований или методов обработки естественного языка осуществляется поиск и извлечение релевантной информации из внешних баз данных, документов или других источников;

- генерация данных: на втором этапе извлеченная информация используется для дополнения синтетических данных. Модели МО могут быть использованы для интеграции новой информации в существующие синтетические данные.

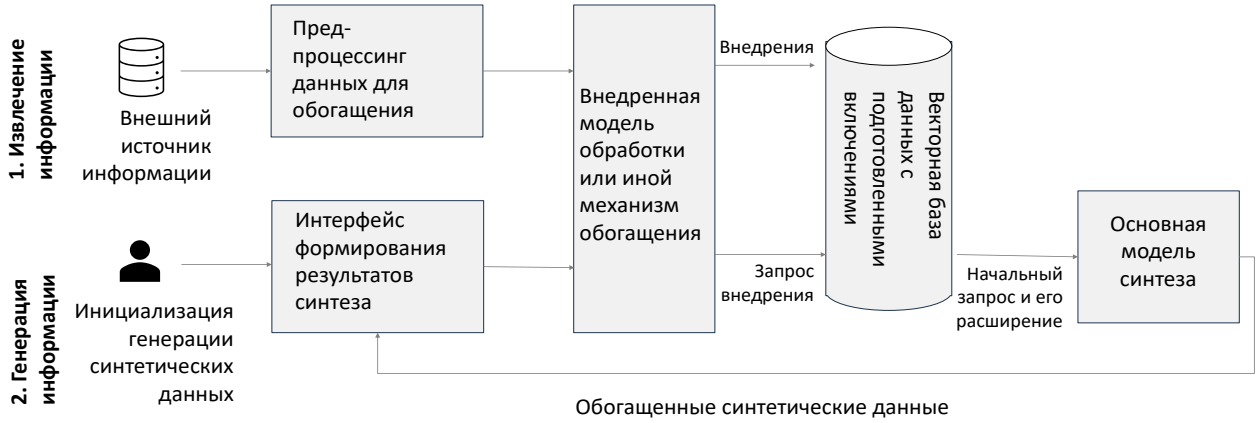


Рисунок В.6 — Принципиальная архитектура синтеза с обогащением

Применение такой архитектуры повышает качество синтеза данных за счет использования дополнительных источников информации, однако требует дополнительных ресурсов для хранения и обработки внешней информации с использованием таких решений, как векторные базы данных.

В.4.2 Семантическое обогащение

Семантическое обогащение— это процесс добавления семантической информации к синтетическим данным для улучшения их интерпретируемости и полезности. Семантическое обогащение включает следующие компоненты:

- аннотация данных: добавление метаданных для описания содержимого и структуры;
- использование онтологий: структурирование данных с помощью онтологий и семантических сетей;
- семантическая интеграция: объединение данных из разных источников с учетом их смысла для повышения согласованности.

Использование семантического обогащения предоставляет следующие преимущества:

- повышение интерпретируемости: семантическое обогащение делает данные более понятными и доступными для пользователей;
- улучшение качества анализа: обогащенные семантической информацией данные обеспечивают более точные и детализированные результаты анализа;
- унификация данных: семантическая интеграция позволяет объединять данные из различных источников, что способствует их согласованности и целостности.

Библиография

[1] ИСО 8000-61:2016 Качество данных. Часть 61. Управление качеством данных. Базовая модель процесса (Data Quality — Part 61: Data Quality Management: Process Reference Model)

[2] ISO/IEC/IEEE 42030:2019 Программное обеспечение, системы и корпоративная среда. Схема оценки архитектуры (Software, Systems and Enterprise — Architecture Evaluation Framework)

[3] ИСО 29101:2018 Информационная технология. Методы и средства обеспечения безопасности. Базовая архитектура конфиденциальности (Information Technology — Security Techniques — Privacy Architecture Framework)

[4] ИСО/МЭК 20547-4:2020 Информационная технология. Эталонная архитектура больших данных. Часть 4. Безопасность и приватность (Information Technology — Big Data Reference Architecture — Part 4: Security and Privacy)

[5] ИСО 19115-1:2014 Географическая информация. Метаданные. Часть 1. Основные положения (Geographic information — Metadata — Part 1: Fundamentals)

[6] ИСО/МЭК 20889:2018 Терминология и классификация методов деидентификации данных повышенной конфиденциальности (Privacy enhancing data de-identification terminology and classification of techniques)

[7] NIST.SP.800-226 Guidelines for Evaluating Differential, 2025

[8] ИСО/МЭК 5259-3:2024 Искусственный интеллект. Качество данных для аналитики и машинного обучения (ML). Часть 3. Требования и руководство по менеджменту качества данных [Artificial Intelligence — Data Quality for Analytics and Machine Learning (ML) — Part 3: Data Quality Management Requirements and Guidelines]

УДК 004.89:004.9:519.68:006.354 ОКС 35.020

Ключевые слова: синтез данных, архитектура процесса синтеза данных, методы синтеза

Руководитель организации-разработчика

Ассоциация Больших Данных
наименование организации

Исполнительный директор	_____	**А.В. Нейман**
должность	личная подпись	инициалы, фамилия

Руководитель разработки

_____	_____	
должность	личная подпись	инициалы, фамилия

Исполнитель:

_____	_____	
должность	личная подпись	инициалы, фамилия