DOI 10.37882/2223-2966.2025.06.51

ИСПОЛЬЗОВАНИЕ ИИ ДЛЯ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ ПРИ РАЗРАБОТКЕ ИС

USE OF AI FOR PERSONAL DATA DEPARTMENT IN IS DEVELOPMENT

G. Shevtsova
K. Dobrinov

Summary. The article discusses approaches to personal data depersonalization using artificial intelligence (AI) technologies in the process of development and maintenance of information systems (IS). Particular attention is paid to the architecture of AI solutions embedded in engineering practices, an overview of modern tools such as autoencoders and differential privacy, as well as examples from the financial and medical sectors is presented. The advantages of intelligent methods over traditional depersonalization schemes are substantiated, with an emphasis on the practical integration of such solutions into the life cycle of an information system.

Keywords: artificial intelligence, data depersonalization, information systems development, machine learning, autoencoders, testing, differential privacy, data engineering.

Шевцова Галина Александровна
Кандидат исторических наук, доцент, Российский государственный гуманитарный университет
shevtsova-g@mail.ru
Добринов Кирилл Павлович
Аспирант, Российский государственный гуманитарный университет
kpdobrinov@gmail.com

Аннотация. В статье рассматриваются подходы к обезличиванию персональных данных с использованием технологий искусственного интеллекта (ИИ) в процессе разработки и сопровождения информационных систем (ИС). Особое внимание уделено архитектуре ИИ-решений, встроенных в инженерные практики, представлен обзор современных инструментов, таких как автоэнкодеры и дифференциальная приватность, а также примеры из финансового и медицинского секторов. Обоснованы преимущества интеллектуальных методов по сравнению с традиционными схемами обезличивания, акцент сделан на практическую интеграцию таких решений в жизненный цикл ИС.

Ключевые слова: искусственный интеллект, обезличивание данных, разработка информационных систем, машинное обучение, автоэнкодеры, тестирование, дифференциальная приватность, инженерия данных.

С каждым годом персональные данные всё активнее вовлекаются в работу технологичных продуктов и обогащают цифровой след. Банковские операции, медицинские записи, действия в социальных сетях — всё это генерирует огромные массивы информации, которые собираются, хранятся и обрабатываются различными информационными системами. На этом фоне задача защиты персональных данных становится вопросом безопасности, правового соответствия и этики. При этом такие данные применяются в том числе и на этапах проектирования, тестирования и аналитики информационных систем, а также при обучении алгоритмов машинного обучения. На каждом из этих этапов сохраняется риск утечки или неправомерного доступа, особенно если в работе используются чувствительные пользовательские данные.

Обезличивание — один из базовых методов защиты информации, позволяющий исключить возможность установления личности субъекта данных. Однако традиционные подходы, включая удаление идентификаторов, генерализацию и маскирование, обладают существенными ограничениями. Они слабо масштабируются, приводят к потере значимой информации и не гарантируют защиту от повторной идентификации при наличии сторонних источников [1].

В качестве альтернативы всё чаще рассматриваются решения, основанные на технологиях искусственного интеллекта. Модели машинного обучения способны автоматически обнаруживать чувствительные признаки, учитывать взаимосвязи между полями и проводить обезличивание с сохранением структурной и аналитической целостности данных. Это особенно актуально в процессе разработки информационных систем, где важны как безопасность, так и достоверность рабочих наборов данных.

Настоящая статья посвящена анализу подходов к обезличиванию персональных данных с применением ИИ в рамках жизненного цикла разработки и сопровождения информационных систем.

Обезличивание персональных данных традиционно реализуется с использованием формальных подходов, направленных на исключение прямой связи между информацией и конкретным человеком. К числу таких методов относятся:

- удаление прямых идентификаторов, таких как фамилия, имя, отчество, номер паспорта, e-mail;
- генерализация, при которой точные значения заменяются обобщёнными категориями (например, возраст → диапазон, адрес → город);

— псевдонимизация — подмена идентификаторов случайными или уникальными кодами;
— маскирование, то есть частичное скрытие значений (например, отображение только последних цифр номера телефона);
— добавление статистического шума к числовым или категориальным данным.

Эти методы применяются в нормативной практике, особенно в задачах, где важно быстро исключить идентификаторы при передаче данных внешним системам или при публикации агрегированной информации. Они просты в реализации и не требуют вычислительно затратных операций.

Однако несмотря на широкое распространение, данные подходы имеют целый ряд ограничений. Во-первых, они часто нарушают внутреннюю структуру данных, делая их малопригодными для анализа. Во-вторых, они плохо адаптируются к разнообразным форматам и взаимосвязанным атрибутам, что особенно критично для современных информационных систем. В-третьих, их недостаточная устойчивость к повторной идентификации при использовании внешних источников давно зафиксирована в научной и правовой литературе — например, в документах Европейской комиссии по применению GDPR [2], а также в российских исследованиях, посвящённых оценке уязвимостей классических методов [1].

Кроме того, такие методы часто применяются вне контекста бизнес-процессов или архитектуры ИС. Это создаёт риск того, что обезличивание рассматривается как вспомогательная операция, а не как встроенный механизм обеспечения приватности на уровне инженерии. В результате решения оказываются негибкими и не соответствуют требованиям к интеграции, масштабируемости и повторному использованию.

В ответ на ограничения традиционных подходов всё большую актуальность приобретают методы обезличивания, основанные на технологиях искусственного интеллекта (ИИ). Такие методы ориентированы не на механическое удаление или подмену полей, а на контекстный анализ структуры данных, выявление чувствительных признаков и трансформацию информации с учётом её внутренней логики. Это особенно важно в условиях сложных, многоуровневых или слабо структурированных данных, которые типичны для современных информационных систем.

Ключевым преимуществом ИИ является способность обнаруживать неявные зависимости между признаками, что невозможно при использовании статичных или заранее заданных схем обезличивания. Например, если фамилия удалена, но остаются профессия, возраст и регион, алгоритмы машинного обучения могут определить, подлежит ли запись дальнейшему обезличиванию, даже если атрибут формально не является идентифицирующим.

В числе наиболее эффективных ИИ-подходов:
— автоэнкодеры — нейросетевые модели, способные формировать обобщённое представление данных, устраняя индивидуальные особенности и позволяя восстанавливать их без идентификационных признаков [3];
— обучаемые классификаторы — определяют уровень чувствительности конкретных полей и применяют к ним соответствующие стратегии обработки;
— модели на основе дифференциальной приватности — позволяют добавлять шум к результатам запросов или наборам данных, ограничивая вероятность восстановления индивидуальной записи даже при доступе к агрегированной информации [4].

Эти подходы обеспечивают более высокую устойчивость к повторной идентификации, сохраняя при этом возможность использования данных для аналитических, исследовательских и инженерных целей. Особенно ценно это в условиях разработки ИС, где важно тестировать системы на реалистичных наборах, но без риска утечки. Кроме того, ИИ-системы можно адаптировать под конкретную предметную область, обучив модель на частотных паттернах и внутренних правилах конкретного бизнеса. Это делает возможным не просто «обезличивание по шаблону», а интеллектуальную трансформацию данных с учётом логики их применения в рамках информационной системы.

Эффективное применение методов ИИ для обезличивания данных требует не только выбора подходящей модели, но и построения целостной архитектуры, которая может быть встроена в общий контур разработки информационной системы. Это особенно важно, если обезличивание выполняется не как разовая процедура, а как непрерывный процесс, сопровождающий разработку, тестирование, аналитику и поддержку продукта.

В условиях промышленной разработки архитектура таких систем должна учитывать требования к масштабируемости, воспроизводимости, интеграции с другими сервисами, а также соответствие нормативным и внутренним политикам безопасности. Простейшая схема «загрузил — обезличил — выгрузил» не подходит для рабочих ИС, где данные меняются постоянно, работают CI/CD-пайплайны, а разные команды обмениваются информацией через API.

В этой связи архитектура системы обезличивания на основе ИИ, как правило, строится по модульному принципу и включает в себя следующие ключевые компоненты:

1. Модуль предобработки данных

На первом этапе происходит подготовка данных для анализа и последующего обезличивания. Этот модуль отвечает за очистку (удаление дубликатов, заполнение пропусков), нормализацию (приведение значений к стандартным форматам) и извлечение метаинформации о типах признаков. Именно здесь формируется «база» для корректной работы моделей — некорректно размеченные или загрязнённые данные могут привести к тому, что ИИ ошибочно интерпретирует чувствительные поля.

Подготовленные данные также подвергаются первичной агрегации: выделяются типовые шаблоны и статистические распределения, которые используются далее для калибровки моделей обезличивания.

2. Модуль классификации и разметки чувствительных данных

Ключевая задача этого модуля — определить, какие именно поля содержат персональную информацию и требуют особого обращения. Здесь применяются обученные модели (например, градиентный бустинг или логистическая регрессия), а также правила, основанные на регулярных выражениях и списках атрибутов (PII, SPII и др.).

В продвинутых решениях используется комбинация автоматической классификации и ручной верификации. Такой подход позволяет добиться баланса между масштабируемостью и точностью, особенно в системах, работающих с нестандартными или локализованными данными.

3. Ядро обезличивания (ИИ-модуль)

Это центральная часть системы. В зависимости от задач здесь применяются разные алгоритмы:

— Автоэнкодеры — нейросети, которые обучаются сжимать данные в латентное пространство и восстанавливать их обратно, исключая при этом идентифицирующую информацию [3].
— Генеративные модели (GAN, VAE) — используются для создания синтетических данных, статистически схожих с оригинальными.
— Алгоритмы на основе дифференциальной приватности — добавляют контролируемый шум, минимизируя вероятность идентификации конкретной записи [4].

Ядро может работать как в пакетном режиме (например, при регулярной выгрузке данных), так и в потоковом — например, при интеграции с веб-приложением или API.

4. Модуль контроля качества и валидации

Обезличивание должно быть не просто выполнено, но и верифицировано. Для этого используются формальные метрики:

— k-анонимность — показывает, сколько записей выглядят идентично в обезличенном наборе;
— l-разнообразие — оценивает, насколько разнообразны чувствительные значения в пределах одинаковых групп;
— t-близость — сравнивает распределения чувствительных признаков до и после обезличивания.

Также важен встроенный аудит: система логгирует, какие данные и каким методом были обработаны, какие версии модели применялись, какие поля подлежали контролю. Это позволяет как соблюсти требования законодательства (например, GDPR и ФЗ-152), так и обеспечить воспроизводимость решений внутри команды.

5. Интерфейс интеграции с другими системами

Практически ни одна система не работает изолированно. Архитектура обезличивания должна предусматривать возможность интеграции с:

— корпоративными базами данных (SQL/NoSQL);
— пайплайнами CI/CD (например, GitLab или Jenkins);
— облачными хранилищами (например, Yandex Cloud);
— API для внешнего и внутреннего обмена данными.

Наличие гибких интерфейсов делает обезличивание частью инфраструктуры, а не внешним инструментом.

Примеры платформ, реализующих такие архитектурные принципы:

— Google Cloud Data Loss Prevention — сервис, предоставляющий API для автоматического обнаружения и маскировки PII в потоках данных [5];
— IBM Watson Knowledge Catalog — комплексная система управления метаданными, обеспечивающая обнаружение, классификацию и защиту данных [6];
— Amazon Macie — инструмент, использующий ML для анализа и защиты данных в облаке AWS, с поддержкой политики управления доступом [7].

Проблема защиты персональных данных особенно остро стоит в условиях активной разработки, внедрения и сопровождения информационных систем. На практике это означает, что данные подвергаются обработке на всех этапах жизненного цикла ИС.

Во многих организациях разработчики, тестировщики и аналитики работают с копиями «боевых» данных, либо в контексте прототипирования, либо для отладки

бизнес-логики. Использование реальных персональных данных в этих целях может приводить к непреднамеренным утечкам, особенно в распределенных командах или при передаче данных подрядчикам. Поэтому обезличивание должно быть встроено в саму культуру разработки — как неотъемлемый элемент процесса, а не внешний

5.1. ИИ-обезличивание в процессах разработки информационных систем

Применение ИИ позволяет автоматизировать обезличивание в ИТ-практиках, обеспечивая:
— сохранение логической структуры и связей между полями (например, имя ↔ пол ↔ регион);
— генерацию реалистичных, но не раскрывающих личность данных;
— обработку потоков данных в рамках CI/CD и DevOps-пайплайнов.

Такие решения востребованы в крупных компаниях, где ценится стабильность процессов и соблюдение требований законодательства (GDPR, ФЗ-152). Например, используются инструменты:
— Tonic.ai, DataMasque — платформы генерации обезличенных данных для разработки и тестирования;
— Google Cloud DLP API — API для автоматического обнаружения и маскировки чувствительных данных [5];
— Faker, Mockaroo — библиотеки для генерации синтетических данных.

Интеграция этих решений в процесс разработки позволяет не только снизить риски, но и ускорить выход продукта за счет стандартизации работы с данными.

5.2. Финансовый сектор

Финансовые учреждения — банки, страховые компании, платёжные сервисы — ежедневно обрабатывают миллионы записей с чувствительной информацией: от транзакций до профилей клиентов. В то же время аналитические и рекомендательные системы требуют доступа к поведенческим паттернам, которые можно использовать без раскрытия личности.

В таких случаях ИИ-обезличивание применяется для маскировки идентификаторов при сохранении ценности данных для анализа. Например, в Tinkoff реализованы решения, позволяющие анонимизировать поведение клиента при тестировании новых функций и построении ML-моделей [8]. Модели сохраняют динамику операций, категории трат и сценарии поведения, но исключают идентифицирующие данные.

Такой подход обеспечивает соблюдение требований закона и внутренней политики безопасности без ущерба для продуктовой аналитики.

5.3. Здравоохранение и медицинские ИС

В медицинской сфере обезличивание данных особенно критично. Электронные медицинские карты, снимки, лабораторные показатели — всё это содержит персональную и биометрическую информацию. Однако такие данные необходимы для научных исследований, управления ресурсами и разработки ИИ-моделей диагностики.

ИИ-обезличивание применяется для:
— удаления лицевых признаков с МРТ и КТ с помощью сверточных нейросетей;
— генерации синтетических данных на основе реальных записей пациентов;
— построения обезличенных выборок для обучения моделей без доступа к оригиналам.

В частности, в клиниках Mayo Clinic применяются гибридные решения, где ИИ-модель удаляет идентификаторы, а эксперты контролируют медицинскую точность [9]. Это обеспечивает баланс между научной пользой и правовой безопасностью.

Примеры из банковской и медицинской отраслей подтверждают, что интеграция ИИ-обезличивания в инженерные процессы позволяет сочетать юридическую корректность, гибкость разработки и аналитическую ценность данных.

В условиях усложнения информационных систем и роста требований к защите персональных данных, обезличивание становится неотъемлемой частью жизненного цикла разработки. Традиционные методы, несмотря на их простоту, всё чаще оказываются недостаточными для обеспечения реальной приватности. Они не учитывают взаимосвязей между данными, не адаптируются к контексту и могут разрушать аналитическую ценность информации. Использование искусственного интеллекта позволяет преодолеть эти ограничения: ИИ-решения способны анализировать структуру данных, автоматически выявлять чувствительные признаки, адаптироваться к предметной области и выполнять обезличивание с минимальными потерями. Это особенно важно в процессе разработки ИС, где данные используются многократно: при тестировании, обучении моделей, интеграции и анализе. Примеры из финансовой и медицинской практики показывают, что интеллектуальные подходы уже успешно применяются в реальных проектах. ИИ-инструменты интегрируются в пайплайны разработки, повышая как безопасность, так и эффективность процессов. Архитектуры таких систем включают в себя модули класси-

фикации, предобработки, контролируемой генерации и валидации — всё это делает обезличивание частью инженерной практики.

Таким образом, ИИ-обезличивание — это не только способ соблюдения требований закона, но и технологически зрелый инструмент, позволяющий разрабатывать информационные системы более гибко, безопасно и масштабируемо. В будущем можно ожидать дальнейшего распространения таких решений, их стандартизации, а также более глубокой интеграции с аналитикой, машинным обучением и безопасной архитектурой хранения данных.

ЛИТЕРАТУРА

1. Алгоритмы реализации методов обезличивания персональных данных в распределённых информационных системах [Электронный ресурс] // КиберЛенинка. — URL: https://cyberleninka.ru/article/n/algoritmy-realizatsii-metodov-obezlichivaniya-personalnyh-dannyh-v-raspredelennyh-informatsionnyh-sistemah (дата обращения: 18.05.2025).
2. General Data Protection Regulation (GDPR) [Электронный ресурс] // Официальный портал Европейской комиссии. — URL: https://gdpr.eu/ (дата обращения: 18.05.2025).
3. Автоэнкодеры: типы архитектур и применение [Электронный ресурс] // Neurohive. — URL: https://neurohive.io/ru/osnovy-data-science/avtojenkoder-tipy-arhitektur-i-primenenie/ (дата обращения: 18.05.2025).
4. Differential Privacy: A Primer for a Non-technical Audience [Электронный ресурс] // OpenMined. — URL: https://blog.openmined.org/differential-privacy-a-primer-for-a-non-technical-audience/ (дата обращения: 18.05.2025).
5. Google Cloud Data Loss Prevention [Электронный ресурс]. — URL: https://cloud.google.com/dlp (дата обращения: 18.05.2025).
6. IBM Watson Knowledge Catalog [Электронный ресурс]. — URL: https://www.ibm.com/cloud/watson-knowledge-catalog (дата обращения: 18.05.2025).
7. Amazon Macie [Электронный ресурс]. — URL: https://aws.amazon.com/macie/ (дата обращения: 18.05.2025).
8. Парадокс ИИ и конфиденциальности данных, а также будущее совместной работы с личными данными [Электронный ресурс] // AppsFlyer. — URL: https://www.appsflyer.com/ru/blog/ceo/ai-privacy-data-collaboration/ (дата обращения: 18.05.2025).
9. AI in Healthcare: Using Artificial Intelligence for Patient Privacy [Электронный ресурс] // Mayo Clinic Platform. — URL: https://platform.mayoclinic.org/articles/ai-and-healthcare-privacy (дата обращения: 18.05.2025).

Журнал «Современная наука: актуальные проблемы теории и практики»