

УДК 004.056:004.421.6

Повышение конфиденциальности в образовательных данных: разработка алгоритма обезличивания с использованием AES/GCM и методов перемешивания

Е. В. Зверева¹, Р. А. Ещенко², Е. С. Добросельская², О. И. Чуйко³, Я. Д. Новак³

¹ Петербургский государственный университет путей сообщения Императора Александра I, Россия, 190031, Санкт-Петербург, Московский пр., 9

² Дальневосточный государственный университет путей сообщения, Россия, 680021, Хабаровск, ул. Серышева, 47

³ Хабаровский институт инфокоммуникаций — филиал Сибирского государственного университета телекоммуникаций и информатики, Россия, 680013, Хабаровск, ул. Ленина, 73

Для цитирования: Зверева Е. В., Ещенко Р. А., Добросельская Е. С., Чуйко О. И., Новак Я. Д. Повышение конфиденциальности в образовательных данных: разработка алгоритма обезличивания с использованием AES/GCM и методов перемешивания // Бюллетень результатов научных исследований. 2026. Вып. 3. С. 165–187. DOI: 10.20295/2223-9987-2026-3-165-187

Аннотация

Цель: проанализировать существующие методы и способы обезличивания персональных данных студентов, а также разработать алгоритмы их обезличивания для обеспечения информационной безопасности в образовательных организациях с учетом требований российского законодательства.

Методы: сравнение существующих подходов для обезличивания персональных данных. Предложен гибридный алгоритм обезличивания персональных данных, который сочетает в себе метод изменения состава или семантики и метод перемешивания, основанный на современных криптографических стандартах. Разработан комплексный алгоритм обезличивания персональных данных. **Результаты:** предоставление организациям образовательной сферы возможности осуществлять легальную передачу сведений в целях статистического анализа, валидации методик, тестирования программных средств или проведения прикладных исследований без нарушения прав субъектов данных, одновременно снижая риски несанкционированного доступа. **Практическая значимость:** создан и экспериментально обоснован комплексный алгоритм для деидентификации персональных данных, предназначенных для применения в инфраструктуре образовательных учреждений. Данный алгоритм, базирующийся на комбинации трансформации семантической структуры, перемешивания записей и использования актуальных криптографических протоколов и стандартов, обеспечивает возможность легитимной обработки персональных данных обучающихся с соблюдением нормативных требований, которые закреплены в отечественном законодательстве об информации, информационной безопасности и персональных данных.

Ключевые слова: обезличивание персональных данных, анонимизация, информационная безопасность, криптография, AES/GCM, PBKDF2, HMAC, SHA-256, Base64, метод перемешивания, метод изменения состава или семантики

Введение

В современном цифровом обществе вопросы обеспечения информационной безопасности приобретают особую актуальность не только в крупных коммерческих организациях или государственных структурах, но и в образовательных учреждениях. Активное внедрение информационных систем в образовательный

процесс, постоянный обмен разными данными о студентах, преподавателях и иных участниках образовательных отношений приводят к необходимости надежной защиты персональных данных (ПД). Университеты располагают массивами различной чувствительной информации: паспортными и налоговыми данными студентов, сведениями о месте жительства, номерами телефонов, результатами обучения и другими персональными сведениями. В этих условиях одним из основных требований становится соответствие используемых методик обработки законодательству РФ.

Федеральный закон от 27 июня 2006 года № 152-ФЗ «О персональных данных» устанавливает права физических лиц, являющихся субъектами персональных данных, ответственность за нарушение процедур обработки, выделяет отдельные принципы хранения, актуализации и уничтожения ПД, а также требует регулярно внедрения организационных и технических мер, гарантирующих анонимность и невозможность несанкционированного раскрытия информации о субъектах.

Для университетской среды задачи защиты ПД студентов осложняются дополнительными условиями — высокой динамикой контингента обучающихся, разнообразием используемых информационных ресурсов, необходимостью соблюдения как требований информационной безопасности, так и доступности данных для учебных, исследовательских и административных целей. Решение вопросов анонимизации данных не может быть сведено к простой задаче удаления идентификаторов или маскировки отдельных полей. В практических условиях университеты сталкиваются с ситуациями, когда обезличенные данные должны быть пригодны для статистических исследований, внутреннего аудита, формирования отчетности, но при этом не должны приводить к возможности идентификации реального лица без специально регулируемого доступа.

Современные подходы к обезличиванию данных предлагают широкий инструментарий методов. Среди них можно выделить введение идентификаторов, изменение состава или семантики, декомпозицию, а также перемешивание. Каждый из методов обладает собственными преимуществами и ограничениями, а их комбинация позволяет выстроить индивидуальный контур безопасности с учетом специфики обрабатываемых данных и задач образовательного учреждения. Особое значение имеет качество алгоритмической реализации: решения обязаны быть не только устойчивыми к атакам на приватность, но и не снижать полезность данных для легитимных целей поиска, статистики и администрирования.

В рамках научного исследования поставлена задача разработки комплексного алгоритма обезличивания персональных данных студентов университета для обеспечения информационной безопасности. Цели настоящей работы: анализ предметной области, классификация и сравнение существующих методов обезличивания, проектирование и программная реализация алгоритма обезличивания.

Обзор нормативно-правовой базы

Основным нормативно-правовым актом является федеральный закон от 27 июня 2006 года № 152-ФЗ «О персональных данных», который регулирует отношения, связанные с обработкой персональных данных с использованием средств автоматизации [1]. Положения закона закрепляют:

- принципы и условия обработки персональных данных;
- права субъекта персональных данных;
- обязанности оператора;
- контроль за обработкой персональных данных;
- ответственность за нарушение данного закона.

Контроль и надзор за соответствием обработки персональных данных требованиям законодательства Российской Федерации в области персональных данных осуществляется Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор) [1, 5].

Состав и содержание требований к защите персональных данных при их обработке в информационных системах устанавливаются Федеральной службой безопасности (ФСБ) и Федеральной службой технического экспорта и контроля (ФСТЭК) [1, 3, 4].

Помимо федерального закона № 152-ФЗ, обезличивание персональных данных закрепляется постановлением Правительства Российской Федерации от 21 марта 2012 года № 211 «Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных» и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами» [6].

В приказе Роскомнадзора от 5 сентября 2013 года № 996 «Об утверждении требований и методов по обезличиванию персональных данных» закреплены требования и методы по обезличиванию персональных данных, обрабатываемых в информационных системах [9]. Перечень методов, закрепленных данным документом:

- введение идентификаторов;
- изменение состава или семантики;
- декомпозиция;
- перемешивание.

Анализ существующих методов обезличивания

Анализ существующих методов обезличивания проводится в соответствии со свойствами обезличенных данных и методов обезличивания [10].

Метод введения идентификаторов

1. Полнота: в таблицу соответствия заносится вся информация, позволяющая идентифицировать субъекта персональных данных.

2. Структурированность: каждому идентификатору однозначно соответствует свой набор данных.

3. Релевантность: отсутствует, так как вид представления персональных данных в запросе и ответе на запрос может отличаться.

4. Семантическая целостность: персональные данные переносятся в таблицу соответствия, при этом их вид представления остается неизменным.

5. Применимость: возможна обработка каждой записи и обезличенных данных в целом без деобезличивания.

6. Анонимность: возможно обеспечить только при определенных правилах выбора идентификаторов и заменяемых ими персональных данных.

Метод изменения состава или семантики

1. Полнота: только если гарантирована сохранность данных при проведении изменений.

2. Структурированность: между значениями атрибутов персональных данных субъекта существует связь, которая остается неизменной.

3. Релевантность: в определенных случаях возможно получить семантическое соответствие поискового запроса и полученного ответа на запрос.

4. Семантическая целостность: только если сохраняется семантика данных при проведении изменений в составе персональных данных.

5. Применимость: позволяет осуществлять обработку, не требующую деобезличивания всего объема данных о субъектах.

6. Анонимность: удаление или обобщение части данных приводит к неоднозначности при идентификации с использованием обезличенных данных.

Метод декомпозиции

1. Полнота: информация о субъектах персональных данных не удаляется, а переносится в другое хранилище.

2. Структурированность: существующая связь между записями в разных хранилищах позволяет однозначно сопоставлять их.

3. Релевантность: возможно получить семантическое соответствие поискового запроса и полученного ответа на запрос.

4. Семантическая целостность: семантика и вид представления данных о субъекте не изменяются.

5. Применимость: позволяет осуществлять обработку данных, расположенных в одном хранилище, как независимо от другого, так и при совместном их использовании, без деобезличивания всего объема обезличенных данных.

6. Анонимность: только при достаточно сложных связях между хранилищами и защите хранилищ от несанкционированного доступа.

Метод перемешивания

1. Полнота: сохраняется вся информация о субъектах персональных данных.

2. Структурированность: при деобезличивании персональных данных связи между ними полностью восстанавливаются.

3. Релевантность: возможно получить семантическое соответствие поискового запроса и полученного ответа на запрос.

4. Семантическая целостность: вид представления данных о субъекте и семантика данных остаются неизменными.

5. Применимость: при наличии доступа к дополнительной информации может осуществляться обработка как отдельных записей о субъектах, так и всех данных без деобезличивания всего объема обезличенных данных.

6. Анонимность: данные перемешиваются по каждому отдельному атрибуту записи о субъекте, что не позволяет без доступа к дополнительной информации определить принадлежность тех или иных данных конкретному субъекту.

По результатам проведенного анализа составлена таблица соответствия методов обезличивания свойствам обезличенных данных и методов обезличивания, где знаком «+» показано безусловное наличие свойства, знаком «*» — наличие свойства при выполнении условия, а знаком «–» — отсутствие свойства.

ТАБЛИЦА. Соответствие методов обезличивания свойствам обезличенных данных

Метод обезличивания Свойства обезличенных данных	Метод введения идентификаторов	Метод изменения состава или семантики	Метод декомпозиции	Метод перемешивания
Полнота	+	*	+	+
Структурированность	+	+	+	+
Релевантность	–	+	+	+
Семантическая целостность	+	*	+	+
Применимость	+	+	+	+
Анонимность	*	+	*	+

Проведенный анализ показал, что все методы в той или иной мере соответствуют свойствам обезличенных данных и методам обезличивания.

Метод перемешивания соответствует всем свойствам безусловно. У остальных одно или два свойства обеспечиваются только при выполнении определенных условий.

В методических рекомендациях Роскомнадзора по применению приказа № 996 прописано, что для улучшения защищенности данных рекомендуется совместно применять несколько методов [10].

Описание предложенного гибридного алгоритма обезличивания

Разрабатываемый алгоритм для обезличивания данных использует следующие средства для защиты информации:

- Переменная окружения — набор значений, которые определяют поведение и настройки операционной системы, а также других программ, работающих в этой среде. Они используются для хранения информации, необходимой различным приложениям, и для настройки системных параметров [18].

- AES (Advanced Encryption Standard) — симметричный алгоритм блочного шифрования, в котором используется один и тот же криптографический ключ для шифрования и для расшифровки данных. Он работает с блоками размера 128 бит и поддерживает ключи размером 128, 192, 256 бит [16].

- GCM (Galois Counter Mode) — режим работы симметричных блочных шифров, принцип работы которого включает использование счетчика для шифрования, применение функции Галуа для создания тега аутентификации, комбинацию блоков шифротекста и кода аутентификации, а также использование ключа хеширования [17].

- PBKDF2 (Password-Based Key Derivation Function 2) — стандарт формирования ключа на основе пароля. Компонентами алгоритма являются псевдослучайная функция, мастер-пароль, соль, количество итераций, желаемая длина ключа [14].

- HMAC (Hash-based Message Authentication Code) — криптографическая конструкция, которая использует хеш-функцию для создания кода аутентификации сообщения [20].

- SHA-256 (Secure Hash Algorithm 256-bit) — криптографический алгоритм однонаправленных хеш-функций с размером выхода 256 бит, основанный на технологии Меркла — Дамгарда с протоколом сжатия Дэвиса — Мейера [19].

- Соль (salt) — строка входных данных, которая добавляется к исходному массиву данных при вычислении хеша. Она используется для усложнения процесса восстановления исходных данных методом перебора [21].

- Nonce (number only used once) — уникальное числовое значение, которое используется только один раз в процессе шифрования или криптографических операций [21].

- Шифротекст — зашифрованные и нечитаемые данные, которые можно прочитать только после их расшифровки с помощью специального ключа шифрования [17].

- Тег (tag) — специальная последовательность бит, которая добавляется к зашифрованному сообщению для проверки его целостности [16].

- Base64 — стандарт кодирования бинарных данных в текстовый формат, использующий 64 различных символа ASCII, в котором исходные бинарные данные разбиваются на блоки по 6 бит, каждый блок кодируется соответствующим символом из алфавита [20].

Перед началом шифрования ищется существующая переменная окружения с секретом. Переменная должна иметь префикс, указывающий на область применения. В нашем случае используется «DB_», потому что алгоритм меняет данные в базе. Допустимы только символы Base64, и переменная должна быть длиной не менее 256 бит. В случае отсутствия данной переменной берется значение секрета по умолчанию.

Далее создается криптографический ключ для AES, в котором формируется строка-вход на основании секрета, соли и строки с текущей датой и временем. После строка-вход переводится в байты.

Затем с помощью PBKDF2 с HMAC генерируется ключ, использующий для генерации SHA-256, соль, 100 тысяч итераций и длину ключа в 32 байта.

Следующим шагом является превращение символов в целое число, которое затем преобразуется в байтовое число. В коде это преобразование выполняется так:

- строка, состоящая из цифр, парсится в целочисленный тип (int или long) с помощью `int.Parse()` или `long.Parse()`;
- полученное число затем преобразуется в массив байтов фиксированной длины с порядком байт bigendian (старший байт записывается первым). Это делается либо через `BitConverter.GetBytes()` с последующим реверсом (если система littleendian), либо вручную через поразрядные сдвиги.

Для деобезличивания производится обратное преобразование. Далее генерируется nonce для AES. После создается AES-GCM-шифратор, основными компонентами которого являются [13]:

- K — секретный ключ (128, 192 или 256 бит);
- IV — вектор инициализации (обычно 96 бит);
- P — открытый текст для шифрования;
- AAD — дополнительные аутентифицируемые данные;
- C — шифротекст;
- T — тег аутентификации (обычно 128 бит).

Процесс работы AES-GCM-шифратора состоит из следующих шагов [13].

Шаг 1. Генерация подключа для операций в поле Галуа — H :

$$H = E_K(0^{128}), \quad (1)$$

где E_K — функция шифрования AES с ключом K ;
 0^{128} — блок из 128 нулевых бит.

Шаг 2. Формирование счетчика с начальным значением J_0 .

При длине IV равной 96 бит:

$$J_0 = IV \parallel 0^{31} \parallel 1, \quad (2)$$

где 0^{31} — 31 нулевой бит;
1 — один бит со значением 1.

При длине IV , не равной 96 бит:

$$J_0 = GHASH_H \left(IV \parallel 0^s \parallel \left[\text{len}(IV) \right]_{64} \right), \quad (3)$$

где $GHASH_H$ — хеш-функция в поле Галуа с ключом H ;

s — количество нулевых бит для дополнения, чтобы длина была кратна 128;

$\left[\text{len}(IV) \right]_{64}$ — длина IV , представленная 64-битным числом.

Шаг 3. Формирование последовательности счетчиков — J_i :

$$J_i = \text{inc}_{32} (J_{i-1}) \text{ для } i = 1, 2, \dots, n, \quad (4)$$

где inc_{32} — операция инкремента младших 32 бит счетчика;

i — порядковый номер в последовательности счетчиков;

n — количество блоков открытого текста.

Шаг 4. Шифрование блоков открытого текста.

Для всех блоков открытого текста, кроме последнего:

$$C_i = P_i \oplus E_K (J_i) \text{ для } i = 1, 2, \dots, n-1, \quad (5)$$

где C_i — результирующий i -й блок шифротекста;

P_i — i -й блок открытого текста.

Для последнего блока, который может быть неполным:

$$C_n = P_n \oplus MSB_{|P_n|} (E_K (J_n)), \quad (6)$$

где C_n — последний блок шифротекста;

P_n — последний блок открытого текста;

MSB — функция, которая берет только заданное количество старших бит;

$|P_n|$ — размер последнего блока открытого текста в битах.

Шаг 5. Подготовка данных для функции GHASH.

Дополненные до размера, кратного 128 битам, дополнительные аутентифицируемые данные — A' :

$$A' = AAD \parallel 0^s. \quad (7)$$

Дополненный до размера, кратного 128 битам, шифротекст — C' :

$$C' = C \parallel 0^t, \quad (8)$$

где t — количество нулевых бит, которые нужно добавить к шифротексту.

Шаг 6. Вычисление промежуточного значения аутентификации — S :

$$S = GHASH_H (A' \parallel C' \parallel \left[\text{len}(AAD) \right]_{64} \parallel \left[\text{len}(C) \right]_{64}), \quad (9)$$

где $\left[\text{len}(AAD) \right]_{64}$ — длина дополнительных аутентифицируемых данных, представленная 64-битным числом;

$\left[\text{len}(C) \right]_{64}$ — длина шифротекста, представленная 64-битным числом.

Шаг 7. Вычисление итогового аутентификационного тега — T :

$$T = MSB_t(S \oplus E_K(J_0)), \quad (10)$$

где t — длина тега аутентификации в битах.

Шаг 8. Вычисление функции $GHASH$:

$$GHASH_H(X_1, X_2, \dots, X_m) = X_1 \cdot H^m \oplus X_2 \cdot H^{m-1} \oplus \dots \oplus X_m \cdot H^1, \quad (11)$$

где X_i — i -й 128-битный блок входных данных;

H_i — H , возведенное в степень i в поле Галуа.

Все операции в $GHASH$ выполняются в конечном поле Галуа — $GF(2^{128})$, определенном с помощью неприводимого полинома — $p(x)$:

$$p(x) = x^{128} + x^7 + x^2 + x + 1. \quad (12)$$

Шифратор шифрует получившееся ранее байтовое число, получая шифротекст и аутентификационный тег. Далее происходит соединение по-прежнему шифротекста и тега, который кодируется в выходную строку вида base64.

Блок-схема разработанного алгоритма обезличивания персональных данных представлена на рис. 1.

Перечисленные выше шаги и средства используются для обезличивания всех видов данных, представленных в статье.

В процессе работы с данными предусмотрены следующие блоки для обработки возможных ошибок:

1. Проверка на null или пустые строки (`string.IsNullOrEmpty`).
2. Проверка длины и допустимых символов (например, для СНИЛС — `snilsFull.Length >= 9`, для ИНН — `inn.Length == 12 && inn.All(char.IsDigit)`).
3. Проверка формата даты через `DateTime.ParseExact` (при генерации маскированных дат).
4. Проверка соответствия шаблону номера телефона с помощью регулярных выражений.
5. При невыполнении условий функция возвращает пустые строки.
6. В методе обезличивания вызовы обернуты в блоки `try-catch`. В случае исключения записывается пустая строка.
7. При деобезличивании вызовы также обернуты в блоки `try-catch` — при ошибке просто пропускается восстановление поля.
8. Для дат используется `DateTime.ParseExact` с `CultureInfo.InvariantCulture`, что гарантирует строгий формат.

В процессе работы алгоритма компоненты для обезличивания хранятся в бинарном виде. При работе с базой данных используется формат JSON.

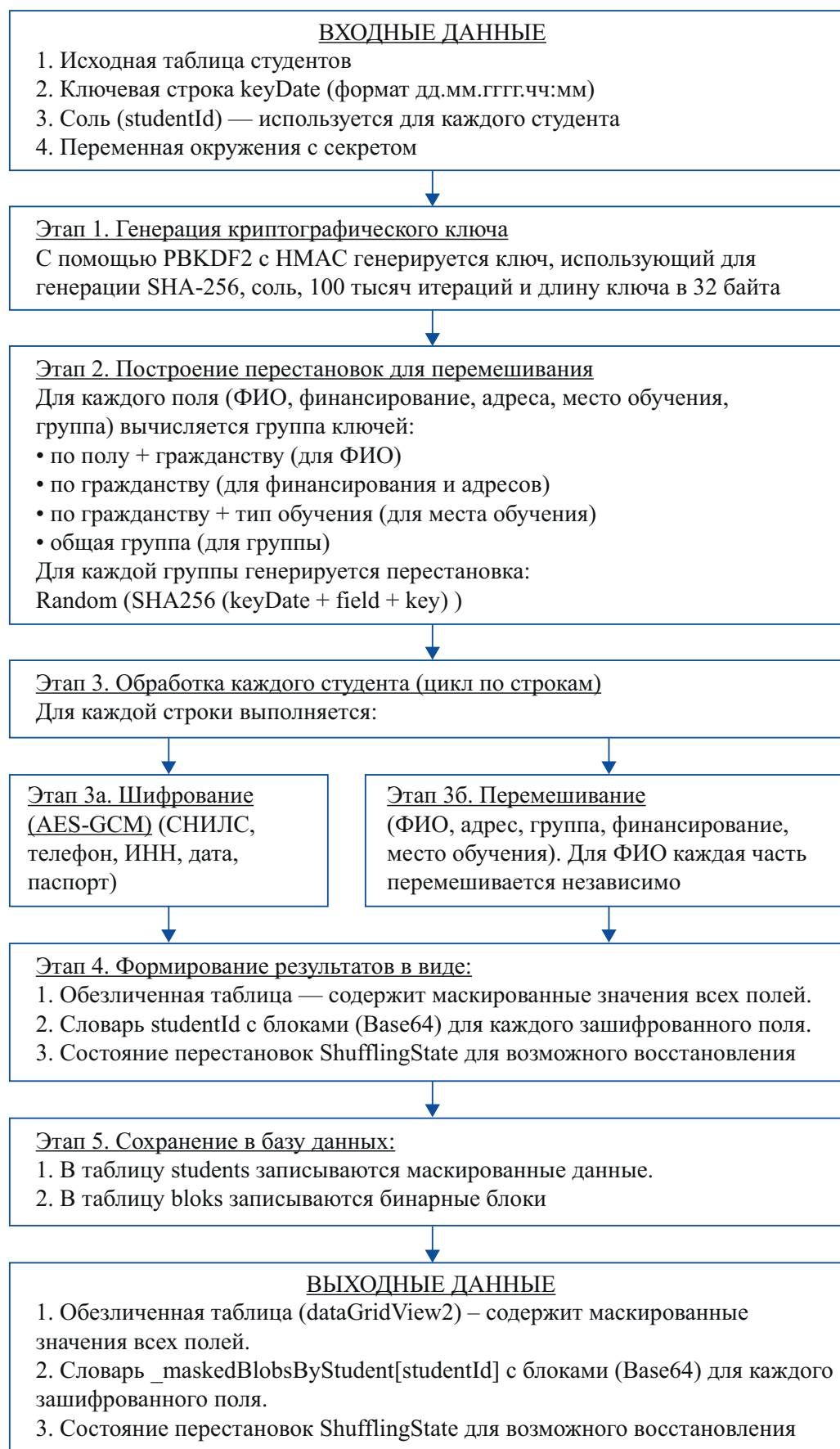


Рис. 1. Блок-схема гибридного алгоритма обезличивания персональных данных

Ниже в пунктах 1–5 представлено конкретизированное использование средств для обезличивания данных по методу изменения состава или семантики, в пункте 6 — по методу перемешивания.

1. Обезличивание СНИЛС

СНИЛС (страховой номер индивидуального лицевого счета) — это уникальный номер человека в системе обязательного пенсионного страхования России. Данный номер присваивается один раз в жизни и не меняется при смене личных данных. СНИЛС состоит из 11 цифр в формате XXX-XXX-XXX-YY, где последние две цифры — контрольное число [2].

Контрольное число СНИЛС рассчитывается следующим образом [11]:

- каждая цифра СНИЛС умножается на номер своей позиции, которые отсчитываются с конца);
- полученные произведения суммируются;
- сумма делится на 101;
- последние две цифры остатка от деления являются контрольным числом.

Для обезличивания СНИЛС берется созданный криптографический ключ, далее генерируется nonce для AES. После первые девять цифр СНИЛС переводятся в число в виде 4 байт со способом хранения big-endian — способ хранения многобайтовых чисел, при котором старший, наиболее значимый байт располагается по наименьшему адресу памяти. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега, которые вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки. Шифротекст также используется еще раз для превращения его в число с ограничением диапазона до девятизначного. По маскированному СНИЛС производится расчет контрольной суммы для проверки валидности номера, который в конце добавляется для получения полного номера СНИЛС, замещающего исходный.

Для деобезличивания производится декодирование base64 в байты. Байты делятся по следующему принципу: nonce — первые 12 байт, далее 4 байта — шифротекст, тег — оставшиеся. Дальше создается AES-GCM-дешифратор с тем же ключом nonce, который с проверкой тега расшифровывает шифротекст. После преобразования получается исходный СНИЛС — первые девять цифр. С помощью вычисления контрольного числа из исходных цифр добавляются последние две цифры. В итоге возвращается полный СНИЛС — строка из 11 цифр.

2. Обезличивание номеров телефонов

Для процесса обезличивания номеров телефонов вначале необходимо создать словарь для разных шаблонов номеров, так как в университете обучаются не только студенты из Российской Федерации. В каждом случае сохраняются коды стран и оператора, но маскируются позиции номера абонента. Например, для номеров телефонов России и Китая будут маскироваться цифры с 5 по 11, так как маска номеров — +7- или 8-YYY-XXX-XX-XX и +86-YY-XXX-XX-XX соответственно,

где первые цифры — код страны, Y — цифры кода оператора, X — цифры абонентского номера. Но для телефонов Монголии будут маскироваться цифры с шестой по одиннадцатую, так как маска номеров — +976-YY-XXX-XXX.

Перед началом обезличивания номер телефона делится на составляющие — префикс и основную часть. Префикс — код страны и оператора, основная часть — номер абонента. Далее эти части проверяются на соответствие словарю.

Для обезличивания номера берется созданный криптографический ключ, далее генерируется nonce для AES. После часть телефона с номером абонента переводится в 8-байтовый массив, который дополняется ведущими нулями при необходимости. Далее байты превращаются в одно число, которое преобразуется в строку, обрезанную под нужную длину символов. Это необходимо для создания маски, состоящей из цифр. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега, которые вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки. После маскированные байты соединяются с кодом оператора для создания виртуального номера с подменой номера абонента.

С целью восстановления оригинального номера извлекается префикс, чтобы знать, сколько цифр было зашифровано. Далее восстанавливается количество цифр, находящихся в основной части номера. Затем производится декодирование base64 в байты. Байты делятся по следующему принципу: nonce — первые 12 байт, далее 8 байтов — шифротекст, тег — оставшиеся. После этого создается AES-GCM-дешифратор с тем же ключом nonce, который с проверкой тега расшифровывает шифротекст. Полученное значение преобразуется в абонентский номер, который затем соединяется с префиксом, чтобы получить исходный номер телефона.

3. Обезличивание номера ИНН

ИНН (идентификационный номер налогоплательщика) — это уникальный номер, который используется налоговыми органами для учета уплаченных и подлежащих уплате налогов и сборов каждого экономического агента [8].

ИНН физического лица состоит из 12 цифр в формате ZZZZXXXXXXYY, где Z — цифры кода подразделения Федеральной налоговой службы России (ФНС России): первые две — код региона согласно Конституции РФ, последние две — код подразделения ФНС России внутри данного региона; X — цифры номера налогоплательщика, уникального в рамках данного подразделения ФНС; Y — цифры проверочного кода согласно коду проверки [8].

Для ИНН из 12 знаков алгоритм проверки выглядит следующим образом [15]:

- вычисляется контрольная сумма по первым одиннадцати знакам со следующими весовыми коэффициентами: 7, 2, 4, 10, 3, 5, 9, 4, 6, 8, 0;
- вычисляется контрольное число $_1$ как остаток от деления контрольной суммы на 11;

- если контрольное число_1 больше девяти, то контрольное число_1 вычисляется как остаток от деления контрольного числа_1 на 10;
- вычисляется контрольная сумма по двенадцати знакам со следующими весовыми коэффициентами: 3, 7, 2, 4, 10, 3, 5, 9, 4, 6, 8, 0;
- вычисляется контрольное число_2 как остаток от деления контрольной суммы на 11;
- если контрольное число_2 больше девяти, то контрольное число_2 вычисляется как остаток от деления контрольного числа_2 на 10;
- контрольное число_1 проверяется с одиннадцатым знаком ИНН и контрольное число_2 проверяется с двенадцатым знаком ИНН. Если полученные результаты равны, то ИНН считается верным.

Данные ИНН обезличиваются в следующем порядке. Сначала проверяется количество цифр, оно должно соответствовать 12. Далее берется созданный криптографический ключ, и генерируется nonce для AES. Затем целое число, состоящее из шести цифр с номером налогоплательщика, переводится в строку из шести цифр, что необходимо для обеспечения количества цифр при подмене реального номера. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега, которые вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки. Далее собирается новый номер ИНН, в котором первые четыре цифры — оригинальные, следующие шесть — зашифрованные, затем вычисляются и подставляются последние две цифры проверочного кода.

Для деобезличивания ИНН сначала проводится проверка, что ИНН — строка из 12 цифр. Затем происходит декодирование base64 в байты. Байты делятся по следующему принципу: nonce — первые 12 байт, далее 3 байта — шифротекст, тег — оставшиеся. Далее создается AES-GCM-дешифратор с тем же ключом nonce, и он с проверкой тега расшифровывает шифротекст. Далее к исходным четырём первым цифрам подставляется получившийся в результате дешифрования оригинальный сегмент. Затем пересчитываются контрольные цифры, которые подставляются в конец.

4. Обезличивание даты

Даты рождения и выдачи паспорта взаимосвязаны друг с другом в случаях первичного получения паспорта в 14 лет, а также смены в связи с достижением 20 и 45 лет. Получение или смена паспорта в таких случаях обычно происходит в течение 90 календарных дней с момента наступления возраста [7]. Следовательно, для корректной замены даты выдачи паспорта необходимо, чтобы ее дата и месяц не превышали 90 дней от даты и месяца рождения.

Предварительно нужно создать проверку на соответствие возрасту на момент выдачи паспорта с максимальной разницей, не превышающей 90 дней. Это делается для проверки информации, что паспорт был выдан по достижении

определенного возраста или по иным причинам: утере, смене фамилии, получении гражданства. Данный этап позволит правильно произвести замену исходных значений даты на маскированное значение.

Для обезличивания даты сначала она разбирается на компоненты. Берется созданный криптографический ключ, далее генерируется nonce для AES. Дата и месяц преобразуются в два байта. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега. Далее байты шифротекста преобразуются в маскированные значения дня и месяца: день — первый байт, для которого делается операция взятия остатка от деления на 31 с прибавлением единицы, месяц — второй байт, для которого делается операция взятия остатка от деления на 12 с прибавлением единицы. Данный шаг необходим, чтобы маскированные дата и месяц были правильного формата. Дальше ведется проверка на соответствие возраста для выдачи паспорта. Если подтверждается, что паспорт получен по достижении возраста, то количество дней между исходными датами прибавляется к маскированной дате рождения, чтобы получить дату выдачи паспорта. Год при этом вычислении не учитывается. Если паспорт был выдан по иным причинам, то дополнительно просто шифруется дата выдачи. Шифротекст, а также аутентификационный тег вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки даты рождения. Для даты выдачи также добавляются параметры nonce и тега аутентификации.

Для деобезличивания маскированные строки дат разбираются на компоненты, берется ключ, используемый для маскирования, декодируются из base64 nonce и тег. Далее происходит восстановление даты рождения. Перебираются значения байтов для дня и месяца с проверкой на соответствие маскированным значениям. При соответствии расшифровываются байты с проверкой тега аутентификации. В случае успешной расшифровки формируются исходные даты и месяцы. Для варианта при получении паспорта, связанном с достижением возраста, вычисляется разница в датах рождения и выдачи паспорта, которая затем складывается с исходной датой рождения, чтобы получить исходное значение даты выдачи паспорта.

Дата документа о предыдущем образовании обезличиваться не будет, так как большое количество людей имеет идентичные даты, что соответствует концепции k-анонимности.

5. Обезличивание серии и номера паспорта

Нумерация паспорта гражданина РФ состоит из серии и номера в формате ZZYY XXXXXX. Серия состоит из 4 цифр. Первые две — номер региона по общероссийскому классификатору территорий муниципальных образований (ОКТМО), выдавшего документ. Вторые две — год, когда был изготовлен бланк паспорта, в нем допускается отклонение до 3 лет. Номер паспорта состоит из шести цифр, они означают номер бланка, на котором печатался паспорт [7].

Буквенно-цифровая нумерация паспорта иностранных граждан состоит из серии и номера в формате YYXXXXXX. Первые два символа — серия паспорта. Она может быть представлена как в виде цифр, так и в виде букв. Оставшиеся семь цифр — номер бланка паспорта [12]. В российских системах внесение информации о таких паспортах производится только в поле номера паспорта, куда записывается буквенно-цифровая нумерация полностью.

Серия паспортов для обоих вариантов будет оставаться без изменения.

Перед началом обезличивания производится анализ на гражданство. Если используются шесть цифр, то считается, что это паспорт гражданина РФ, и все цифры переходят на обезличивание. В случае если используется девять символов, то считается, что это паспорт иностранного гражданина. В таком варианте первые два символа остаются без изменения, а оставшиеся семь подлежат маскировке.

Далее берется созданный криптографический ключ, затем генерируется nonce для AES. После строка, которая подлежит маскировке, преобразуется в число. Затем она кодируется в байты в формат со способом хранения big-endian для шестизначного — в 3 байта, для семизначного — в 4 байта. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега, которые вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки. Зашифрованные байты переводятся обратно в числовое выражение, где делается остаток по модулю 10, чтобы получить числа только нужной длины. В итоге получается маскированное значение паспорта.

Для процесса деобезличивания производится декодирование base64 в байты. Далее по количеству цифр в маскированном номере определяется длина шифротекста. Байты декодирования делятся по следующему принципу: nonce — первые 12 байт, далее 6 или 7 байт — шифротекст, тег — оставшиеся. Дальше создается AES-GCM-дешифратор с тем же ключом nonce, и он с проверкой тега расшифровывает шифротекст. Если в номере было шесть цифр, то просто возвращаются восстановленные цифры, а если семь — то к нему еще прибавляются символы серии.

6. Обезличивание по методу перемешивания

Обезличивание ФИО, места рождения, места прописки и проживания, оконченого учебного заведения, типа обучения, группы происходит по единому принципу, где в качестве входных данных для хэш-функции используется id записи. Также для значений атрибутов курса, квалификации, специальности, профиля берется их связь с группой, в которой обучается студент.

Качественное перемешивание фамилий, имен и отчеств должно сопровождаться привязкой к полу и гражданству. Привязка к полу исключает возможности получить вариант ФИО с несогласованными по роду окончаниями — Петрова Петр Петровна. Привязка к гражданству исключает получение таких сочетаний

фамилии и имени, которые практически невозможны в моноэтнических странах. Например, фамилия и имя Чжан Анастасия вполне возможны в странах на территории СНГ в связи с многообразием этнических групп, но для гражданина Китая — очень редкое сочетание.

Для мест прописки и проживания также берется привязка к гражданству, чтобы не получилось записей, в которых гражданин РФ был бы прописан, например, в Китае.

Оконченное учебное заведение будет привязано к данным о квалификации студента, чтобы обеспечить внесению студентам магистратуры информации об окончании университета, а не школы или техникума.

Перемешивание типа обучения будет производиться только у студентов, чье гражданство включено в список стран СНГ, так как только они могут обучаться на всех трех основах поступления: бюджетная, целевая, договорная.

Корректное перемешивание групп, а вместе с этим и курса, квалификации, специальности, профиля, зависит от года окончания учебного заведения. Группы магистратуры также будут перемешиваться только между студентами, у которых атрибут «квалификация» равен «магистратура». Крайние годы окончания предыдущего образования для каждого курса: 2024 год — 1-й курс, 2023 год — 2-й курс, 2022 год — 3-й курс, 2023 год — 4-й курс, 2025 год — 5-й курс.

У всех описанных данных для обезличивания берется созданный криптографический ключ, далее генерируется nonce для AES. После id переводится в число в виде 4 байт со способом хранения big-endian. AES-GCM-шифратором шифруются байты с получением шифротекста и аутентификационного тега, которые вместе с nonce кодируются в выходную строку вида base64 для хранения и последующей расшифровки. Полученное значение шифротекста используется как ключ для сортировки. При перемешивании происходит перебор по выбранным для привязки в каждом случае значениям. Значения атрибутов затем упорядочиваются в соответствии с ключами сортировки. Конкретно для фамилии, имени, отчества вычисляются разные ключи, в результате чего для каждого атрибута перемешивание происходит независимо.

Для возвращения изначальных данных необходимо сохранение порядка, в котором атрибуты оказались, и на каком месте в перемешанных оказался исходный элемент. Программный интерфейс был выполнен с помощью инструментов Windows Forms.

На рис. 2 представлены две таблицы с исходными и обезличенными данными.

Заключение

Обеспечение информационной безопасности и конфиденциальности персональных данных — задача комплексная, затрагивающая все сферы деятельности современного образовательного учреждения.

The screenshot shows the 'DataAnonym' application window. At the top, there are three tabs: 'Обезличивание' (Anonymization), 'Деобезличивание' (De-anonymization), and 'Логи' (Logs). The 'Обезличивание' tab is active. Below the tabs is a table with 9 columns: ID студента, Фамилия, Имя, Отчество, Пол, Группа, Курс, and Уровень образования. The table contains 6 rows of student data. Below the table, there is a text input field for a date and time in the format 'D01.01.2000T00:00', a button labeled 'Обезличить' (Anonymize), and a 'Сохранить' (Save) button at the bottom right.

ID студента	Фамилия	Имя	Отчество	Пол	Группа	Курс	Уровень образования
1	Ли	Шуан		женский	МО211ИТС	1	Магистратура
2	Кузина	Есения	Демидовна	женский	БО221ИСТ	2	Бакалавриат
3	Воробьева	Елизавета	Романовна	женский	БО221ПИН	2	Бакалавриат
4	Спиридонова	Злата	Егоровна	женский	БО241ПИН	4	Бакалавриат
5	Попова	Мария	Ивановна	женский	СО211КОБ	1	Специалитет
6	Игнатов	Иван	Александрович	мужской	СО211ОПМ	1	Специалитет

Введите дату и время в формате D01.01.2000T00:00: D04.06.2025T15:41

Обезличить

ID студента	Фамилия	Имя	Отчество	Пол	Группа	Курс	Уровень образования
1	Хань	Вен		женский	МО211ИТС	1	Магистратура
2	Кузина	Мария	Ивановна	женский	БО231ИСТ	3	Бакалавриат
3	Попова	Есения	Романовна	женский	БО221ПИН	2	Бакалавриат
4	Воробьева	Елизавета	Егоровна	женский	БО211ИСТ	1	Бакалавриат
5	Спиридонова	Злата	Демидовна	женский	СО211КОБ	1	Специалитет
6	Игнатов	Марк	Глебович	мужской	БО241ИСТ	4	Бакалавриат

Сохранить

Рис. 2. Работа с обезличиванием

Являясь объектами регуляторного контроля и носителями значительных объемов чувствительной информации, университеты нуждаются в современных технологиях и методиках анонимизации, обеспечивающих не только защиту данных, но и их сохранение для образовательных, исследовательских и административных нужд.

В ходе выполнения исследования были проанализированы отечественные нормативно-правовые акты в области обработки персональных данных, описаны их свойства, принципы и порядок работы с ними.

Обезличенные данные должны обладать следующими свойствами: полнота, структурированность, релевантность, семантическая целостность, применимость, анонимность.

В исследовании сравнивались различные методы обезличивания данных: введение идентификаторов, изменение состава или семантики, декомпозиция и перемешивание. У каждого метода есть свои плюсы и минусы, но все они в той или иной степени подходят для обезличивания. Лучше всего требованиям соответствует метод перемешивания — он безоговорочно отвечает всем свойствам. Для большей защиты данных рекомендуется комбинировать несколько методов.

На основе этого анализа был создан алгоритм обезличивания персональных данных — он сочетает в себе два метода: изменение состава или семантики и перемешивание. В основе алгоритма — технологии AES/GCM, PBKDF2, HMAC, SHA256 и Base64.

Алгоритм протестировали на различных персональных данных, включая:

- ФИО, группу, курс и специальность студента;
- дату рождения;
- данные паспорта (серия, номер, дата выдачи);
- информацию об образовании (учебное заведение, дата окончания);
- адреса прописки и проживания;
- номера СНИЛС и ИНН;
- номер телефона.

Список источников

1. О персональных данных: федеральный закон от 27.07.2006 № 152-ФЗ // КонсультантПлюс. URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (дата обращения: 24.03.2026).

2. Об индивидуальном (персонифицированном) учете в системах обязательного пенсионного страхования и обязательного социального страхования: федеральный закон от 01.04.1996 № 27-ФЗ // КонсультантПлюс. URL: https://www.consultant.ru/document/cons_doc_LAW_9839/ (дата обращения: 21.03.2026).

3. Вопросы Федеральной службы безопасности Российской Федерации: указ Президента Российской Федерации от 11.08.2003 № 960 // Гарант. URL: <https://base.garant.ru/12132066/> (дата обращения: 24.03.2026).

4. Вопросы Федеральной службы по техническому и экспортному контролю: указ Президента Российской Федерации от 16.08.2004 № 1085 // КонсультантПлюс. URL: https://www.consultant.ru/document/cons_doc_LAW_14031/ (дата обращения: 24.02.2026).

5. О Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций: постановление Правительства РФ от 16.03.2009 № 228 // КонсультантПлюс: URL: https://www.consultant.ru/document/cons_doc_LAW_85889/ (дата обращения: 24.02.2026).

6. Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных» и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами: постановление Правительства Российской Федерации от 21.03.2012 № 211 // Гарант. URL: <https://base.garant.ru/70152982/> (дата обращения: 24.02.2026).

7. Об утверждении Положения о паспорте гражданина Российской Федерации, образца и описания бланка паспорта гражданина Российской Федерации: постановление Правительства РФ от 23.12.2023 г. № 2267 // КонсультантПлюс: URL: https://www.consultant.ru/document/cons_doc_LAW_466454/ (дата обращения: 21.03.2026).

8. Об утверждении Порядка и условий присвоения, применения, а также изменения идентификационного номера налогоплательщика: приказ ФНС России от 29.06.2012 № ММВ-7-6/435@ // КонсультантПлюс. URL: https://www.consultant.ru/document/cons_doc_LAW_134082/ (дата обращения: 21.03.2026).

9. Об утверждении требований и методов по обезличиванию персональных данных: приказ Роскомнадзора от 05.09.2013 № 996 // Консалтинг.ру. URL: <https://docs.cntd.ru/document/499043850?marker=64U0IK> (дата обращения: 24.02.2026).

10. Методические рекомендации Роскомнадзора по применению приказа Роскомнадзора от 5 сентября 2013 года № 996 «Об утверждении требований и методов по обезличиванию персональных данных» (утв. Роскомнадзором 13.12.2013) // Консалтинг.ру. URL: <https://docs.cntd.ru/document/420281168> (дата обращения: 24.02.2026).

11. Особенности представления страхователями в органы ПФР отчетности, начиная с отчетности с I квартала 2013 года: информационное сообщение Пенсионного фонда Российской Федерации от 21.02.2013 г. // КонсультантПлюс. URL: https://www.consultant.ru/document/cons_doc_LAW_142584/ (дата обращения: 21.03.2026).

12. DOC 9303. Машиносчитываемые проездные документы. Часть 3. Спецификации, общие для всех МСПД (дата введения: 01.01.2015) // ICAO. URL: https://www.icao.int/publications/Documents/9303_p3_cons_ru.pdf (дата обращения: 21.03.2026).

13. ГОСТ Р 34.13–2015. Информационная технология. Криптографическая защита информации. Режимы работы блочных шифров (дата введения: 01.01.2016) // Защита. Госстандарт. URL: <https://protect.gost.ru/v.aspx?control=7&id=200971> (дата обращения: 14.03.2026).

14. Рекомендации по стандартизации Р 1323565.1.040–2022. Информационная технология. Криптографическая защита информации. Парольная защита ключевой информации (дата введения: 01.05.2023) // Защита. Госстандарт. URL: <https://protect.gost.ru/document.aspx?control=7&baseC=101&RegNum=53&DocOnPageCount=500&page=1&id=246142> (дата обращения: 14.03.2026).

15. Алгоритм ключевого ИИН // Кейсистемс. URL: <https://helpbks.keysystems.ru/ro/info/keyinn> (дата обращения: 21.03.2026).

16. Анохин К. Симметричный алгоритм блочного шифрования Advanced Encryption Standard. URL: <https://habr.com/ru/articles/534620/> (дата обращения: 14.03.2026).

17. Бартов Е. Как работает AES (Advanced Encryption Standard). Объяснение для гуманитариев типа меня. URL: <https://vc.ru/dev/656195-kak-rabotaet-aes-advanced-encryption-standard-obyasnenie-dlya-gumanitariyev-tipa-menya> (дата обращения: 14.03.2026).

18. Иванов Д. Переменные окружения для начинающих разработчиков, или Использование .env файла в разработке программного обеспечения. URL: <https://habr.com/ru/companies/gnivc/articles/792082/> (дата обращения: 14.03.2026).

19. Реализация алгоритма SHA-256. URL: <https://habr.com/ru/articles/729260/> (дата обращения: 14.03.2026).

20. Krawczyk H., Eronen P. HMAC-based Extract-and-Expand Key Derivation Function (HKDF) RFC 5869 // Internet Engineering Task Force. 2010. URL: <https://datatracker.ietf.org/doc/html/rfc5869>

Дата поступления: 22.05.2026

Решение о публикации: 08.07.2026

Контактная информация:

ЗВЕРЕВА Елена Валерьевна — кандидат экон. наук, доцент кафедры «Экономика и менеджмент в строительстве»; zverelv@mail.ru

ЕШЕНКО Роман Анатольевич — кандидат техн. наук, доцент кафедры «Цифровая индустрия и защита информации»; era99@yandex.ru

ДОБРОСЕЛЬСКАЯ Екатерина Сергеевна — статистик; katya.dobroselskaya@yandex.ru

ЧУЙКО Ольга Игоревна — кандидат техн. наук, доцент, зав. кафедрой «Информационные технологии»; oskar777@rambler.ru

НОВАК Яна Дмитриевна — преподаватель кафедры «Информационные технологии»; ynovak@bk.ru

Improving Privacy in Educational Data: Developing an Anonymization Algorithm Using AES/GCM and Mixing Methods

E. V. Zvereva¹, R. A. Eshenko², E. S. Dobrosel'skaya², O. I. Chuyko³, Yu. D. Novak³

¹ Emperor Alexander I Petersburg State Transport University, 9 Moskovsky ave., St. Petersburg, 190031, Russia

² Far Eastern State Transport University, 47 Serisheva str., Khabarovsk, 680021, Russia

³ Khabarovsk Institute of Infocommunications, branch of the SSUTI, 73 Lenina str., Khabarovsk, 680013, Russia

For citation: Zvereva E. V., Eshenko R. A., Dobrosel'skaya E. S., Chuyko O. I., Novak Y. D. Improving Privacy in Educational Data: Developing an Anonymization Algorithm Using AES/GCM and Mixing Methods // Bulletin of Scientific Research Results, 2026, iss. 3, pp. 165–187. DOI: 10.20295/2223-9987-2026-3-165-187 (In Russian)

Abstract

Objective: to analyze existing methods and methods of depersonalization of students' personal data, as well as to develop algorithms for their depersonalization to ensure information security in educational institutions, taking into account the requirements of Russian legislation. **Methods:** comparison of existing approaches for depersonalization of personal data. A hybrid algorithm for depersonalizing personal data is proposed, which combines the method of changing the composition or semantics and the mixing method based on modern cryptographic standards. A comprehensive algorithm for depersonalization of personal data has been developed. **Results:** enabling educational organizations to legally transfer information for statistical analysis, validation of methodologies, software testing, or applied research purposes without violating the rights of data subjects, while reducing the risks of unauthorized access. **Practical significance:** a comprehensive algorithm for de-identification of personal data intended for use in the infrastructure of educational institutions has been created and experimentally substantiated. This algorithm, based on a combination of semantic structure transformation, mixing of records and the use of up-to-date cryptographic protocols and standards, provides the possibility of legitimate processing of students' personal data in compliance with regulatory requirements that are enshrined in Russian legislation on information, information security and personal data.

Keywords: depersonalization of personal data, anonymization, information security, cryptography, AES/GCM, PBKDF2, HMAC, SHA-256, Base64, mixing method, a method of changing composition or semantics

References

1. O personal'nykh dannyykh: federal'nyj zakon ot 27.07.2006 № 152-FZ [On Personal Data, Federal Law of July 27, 2006, no. 152-FZ], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (accessed: March 24, 2026). (In Russian)
2. Ob individual'nom (personifitsirovannom) uchete v sistemakh obyazatel'nogo pensionnogo strakhovaniya i obyazatel'nogo sotsial'nogo strakhovaniya: federal'nyj zakon ot 01.04.1996 № 27-FZ [On Individual (Personified) Accounting in the Systems of Mandatory Pension Insurance and Mandatory Social Insurance, Federal Law of April 1, 1996, no. 27-FZ], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_9839/ (accessed: March 21, 2026). (In Russian)
3. Voprosy Federal'noj sluzhby bezopasnosti Rossijskoj Federatsii: ukaz Prezidenta Rossijskoj Federatsii ot 11.08.2003 № 960 [Issues of the Federal Security Service of the Russian Federation, Decree of the President of the Russian Federation of August 11, 2003, no. 960], Garant. URL: <https://base.garant.ru/12132066/> (accessed: March 24, 2026). (In Russian)
4. Voprosy Federal'noj sluzhby po tekhnicheskomu i eksportnomu kontrolyu: ukaz Prezidenta Rossijskoj Federatsii ot 16.08.2004 № 1085 [Issues of the Federal Service for Technical and Export Control, Decree of the President of the Russian Federation of August 16, 2004, no. 1085], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_14031/ (accessed: February 24, 2026). (In Russian)
5. O Federal'noj sluzhbe po nadzoru v sfere svyazi, informatsionnykh tekhnologij i massovykh kommunikatsij: postanovlenie Pravitel'stva RF ot 16.03.2009 № 228 [On the Federal Service for Supervision in the Sphere of Communications, Information Technologies, and Mass Communication, Decree of the Government of the Russian Federation of March 16, 2009, no. 228], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_85889/ (accessed: February 24, 2026). (In Russian)
6. Ob utverzhdenii perechnya mer, napravlennykh na obespechenie vypolneniya obyazannostej, predusmotrennykh Federal'nym zakonom «O personal'nykh dannyykh» i prinyatymi v sootvetstvii s nim normativnymi pravovymi aktami, operatorami, yavlyayushchimisya gosudarstvennymi ili munitsipal'nymi organami: postanovlenie Pravitel'stva Rossijskoj Federatsii ot 21.03.2012 № 211 [On Approval of the List of Measures Aimed at Ensuring Compliance with the Obligations Provided by the Federal Law “On Personal Data” and the Regulatory Legal Acts Adopted in Accordance Therewith by Operators that are State or Municipal Bodies, Decree of the Government of the Russian Federation of March 21, 2012, no. 211], Garant. URL: <https://base.garant.ru/70152982/> (accessed: February 24, 2026). (In Russian)
7. Ob utverzhdenii Polozheniya o pasporte grazhdanina Rossijskoj Federatsii, obraztsa i opisaniya blanka pasporta grazhdanina Rossijskoj Federatsii: postanovlenie Pravitel'stva RF ot 23.12.2023 g. № 2267 [On Approval of the Regulations on the Passport of a Citizen of the Russian Federation, the Model, and the Description of the Passport Blank of a Citizen of the Russian Federation, Decree of the Government of the Russian Federation of December 23, 2023, no. 2267], Consultant Plus, URL: https://www.consultant.ru/document/cons_doc_LAW_466454/ (accessed: March 21, 2026). (In Russian)

8. Ob utverzhdenii Poryadka i uslovij prisvoeniya, primeneniya, a takzhe izmeneniya identifikatsionnogo nomera nalogoplatel'shchika: prikaz FNS Rossii ot 29.06.2012 № MMV-7-6/435@ [On Approval of the Procedure and Conditions for Assigning, Using, and Changing the Taxpayer Identification Number], Order of the Federal Tax Service of Russia of June 29, 2012, no. MMM-7-6/435@], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_134082/ (accessed: March 21, 2026). (In Russian)

9. Ob utverzhdenii trebovanij i metodov po obezlichivaniyu personal'nykh dannykh: prikaz Roskomnadzora ot 05.09.2013 № 996 [On Approval of the Requirements and Methods for Anonymizing Personal Data, Order of Roskomnadzor of September 5, 2013, no. 996], Konsalting.ru. URL: <https://docs.cntd.ru/document/499043850?marker=64U0IK> (accessed: February 24, 2026). (In Russian)

10. Metodicheskie rekomendatsii Roskomnadzora po primeneniyu prikaza Roskomnadzora ot 5 sentyabrya 2013 goda № 996 “Ob utverzhdenii trebovanij i metodov po obezlichivaniyu personal'nykh dannykh”, utv. Roskomnadzorom 13.12.2013 [Methodological Recommendations of Roskomnadzor of December 13, 2013, on the Application of the Order of Roskomnadzor of September 5, 2013, no. 996 “On Approval of the Requirements and Methods for Anonymizing Personal Data”], Konsalting.ru. URL: <https://docs.cntd.ru/document/420281168> (accessed: February 24, 2026). (In Russian)

11. Osobennosti predstavleniya strakhovatelyami v organy PFR otchetnosti, nachinaya s otchetnosti s I kvartala 2013 goda: informatsionnoe soobshchenie Pensionnogo fonda Rossijskoj Federatsii ot 21.02.2013 goda [Features of Reporting by Insurers to the PFR Bodies, Starting with the Reporting for the First Quarter of 2013, Information Message of the Pension Fund of the Russian Federation of February 21, 2013], Consultant Plus. URL: https://www.consultant.ru/document/cons_doc_LAW_142584/ (accessed: March 21, 2026). (In Russian)

12. DOC 9303. Mashinoschityvaemye proezdnye dokumenty. Chast' 3. Spetsifikatsii, obshchie dlya vsekh MSPD, data vvedeniya: 01.01.2015 [DOC 9303. Machine-Readable Travel Documents. Part 3. Specifications Common to All MRTDs, date of introduction: January 1, 2015], ICAO. URL: https://www.icao.int/publications/Documents/9303_p3_cons_ru.pdf (accessed: March 21, 2026). (In Russian)

13. GOST R 34.13-2015. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Rezhimy raboty bloknykh shifrov, data vvedeniya: 01.01.2016 [GOST R 34.13-2015. Information Technology. Cryptographic Protection of Information. Modes of Operation for Block Ciphers], date of introduction: January 1, 2016], Zashchita. Gosstandart [Protection. Gosstandart]. URL: <https://protect.gost.ru/v.aspx?control=7&id=200971> (accessed: March 14, 2026). (In Russian)

14. Rekomendatsii po standartizatsii R 1323565.1.040-2022. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Parol'naya zashchita klyuchevoj informatsii, data vvedeniya: 01.05.2023 [Recommendations for Standardization R 1323565.1.040-2022. Information Technology. Cryptographic Protection of Information. Password Protection of Key Information], date of introduction: May 1, 2023], Zashchita. Gosstandart [Protection. Gosstandart]. URL:

<https://protect.gost.ru/document.aspx?control=7& baseC=101& RegNum=53& DocOnPageCount=500& page=1& id=246142> (accessed: March 14, 2026). (In Russian)

15. Algoritm klyuchevaniya INN [Algorithm for Keying INN], Kejsystems. URL: <https://helpbks.keysystems.ru/ro/info/keyinn> (accessed: March 21, 2026). (In Russian)

16. Anokhin K. Simmetrichnyj algoritm blokovogo shifrovaniya Advanced Encryption Standart [Symmetric Block Encryption Algorithm Advanced Encryption Standard]. URL: <https://habr.com/ru/articles/534620/> (accessed: March 14, 2026). (In Russian)

17. Bartov E. Kak rabotaet AES (Advanced Encryption Standard). Ob"yasnenie dlya gumanitariyev tipa menya [How AES (Advanced Encryption Standard) Works. An Explanation for Humanitarians Like Me]. URL: <https://vc.ru/dev/656195-kak-rabotaet-aes-advanced-encryption-standard-obyasnenie-dlya-gumanitariyev-tipa-menya> (accessed: March 14, 2026). (In Russian)

18. Ivanov D. Peremennye okruzheniya dlya nachinayushchikh razrabotchikov, ili Ispol'zovanie .env fayla v razrabotke programmnogo obespecheniya [Environment Variables for Beginner Developers or Using a .env File in Software Development]. URL: <https://habr.com/ru/companies/gnivc/articles/792082/> (accessed: March 14, 2026). (In Russian)

19. Realizatsiya algoritma SHA-256 [Implementation of the SHA-256 Algorithm]. URL: <https://habr.com/ru/articles/729260/> (accessed: March 14, 2026). (In Russian)

20. Krawczyk H., Eronen P. HMAC-based Extract-and-Expand Key Derivation Function (HKDF) RFC 5869, Internet Engineering Task Force, 2010. URL: <https://datatracker.ietf.org/doc/html/rfc5869>

Received: May 22, 2026

Accepted: July 08, 2026

Author's information:

Elena V. ZVEREVA — PhD in Economics, Associate Professor of the “Economics and Management in Construction” Department; zverelv@mail.ru

Roman A. ESHENKO — PhD in Engineering, Associate Professor of the “Digital Industry and Information Security Department”; era99@yandex.ru

Ekaterina S. DOBROSEL'SKAYA — Statistician; katya.dobroselskaya@yandex.ru

Ol'ga I. CHUYKO — PhD in Engineering, Associate Professor, Head of the “Information Technology” Department; oskar777@rambler.ru

Yana D. NOVAK — Lecturer of the “Information Technology” Department; ynovak@bk.ru