

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 004

DOI: 10.17212/2782-2230-2026-3-64-83

**РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
ДЛЯ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ
В ИСПДн BITRIX24***

Т.А. СИНЮШКИН¹, Г.А. ИСАЕВ²

¹ РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации. E-mail: kaf_zi@corp.nstu.ru

² РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, старший преподаватель кафедры защиты информации. E-mail: g.isaev@corp.nstu.ru

В статье рассматривается разработка программного обеспечения для обезличивания персональных данных в информационной системе персональных данных (ИСПДн) Bitrix24. Актуальность работы обусловлена необходимостью использования данных CRM-систем для тестирования, аналитики и обучения при одновременном соблюдении требований Федерального закона № 152-ФЗ и Постановления Правительства РФ № 1154. Предложен адаптивный подход, основанный на классификации атрибутов CRM-сущностей по четырем категориям: прямые идентификаторы, квазиидентификаторы, чувствительные и служебные атрибуты. Разработан алгоритм динамического выбора методов обезличивания (маскирование, псевдонимизация, обобщение, подавление) в зависимости от категории поля и оцененного риска повторной идентификации. Представлена модульная архитектура программного обеспечения, включающая интеграционный модуль с Bitrix24, модуль обезличивания и ETL-подсистему. Экспериментальная оценка на синтетических данных показала снижение риска повторной идентификации с 70 до 22,5 % при сохранении 100 % логических связей между сущностями.

Ключевые слова: персональные данные, обезличивание, Bitrix24, CRM, ИСПДн, квазиидентификаторы, повторная идентификация, псевдонимизация, маскирование, ETL, API, защита информации

* Статья получена 16 июля 2026 г.

ВВЕДЕНИЕ

Актуальность темы обусловлена тем, что в корпоративной практике часто возникает необходимость использовать данные CRM-систем для тестирования программного обеспечения, построения отчетности, аналитики и обучения сотрудников. При этом применение реальных персональных данных в таких сценариях создает риск их неправомерного раскрытия, утечки или повторной идентификации субъектов. Полное удаление всех значимых сведений также не является оптимальным решением, поскольку в этом случае данные теряют прикладную ценность и не могут использоваться для анализа бизнес-процессов или проверки работы программных систем.

Особую сложность представляет обезличивание данных в CRM-системах, поскольку такие данные имеют связанную структуру. В Bitrix24 информация о контактах, компаниях, сделках, лидах, задачах и комментариях логически взаимосвязана. Персональные данные могут находиться не только в стандартных полях, но и в пользовательских атрибутах, текстовых комментариях и описаниях. Даже после удаления прямых идентификаторов сохраняется риск повторной идентификации субъекта по комбинации квазиидентификаторов.

Проблема исследования заключается в необходимости разработки подхода и программного решения для адаптивного обезличивания персональных данных в системе Bitrix24, обеспечивающего снижение риска повторной идентификации при сохранении структуры и прикладной ценности CRM-данных.

Целью настоящей работы является разработка программного обеспечения для обезличивания персональных данных в системе Bitrix24, обеспечивающего снижение риска повторной идентификации субъектов при сохранении полезности и логической структуры CRM-данных.

1. НОРМАТИВНАЯ БАЗА ОБРАБОТКИ И ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ В ИСПДн

1.1. ПОНЯТИЕ ПЕРСОНАЛЬНЫХ ДАННЫХ И ИХ КЛАССИФИКАЦИЯ

В соответствии с Федеральным законом № 152-ФЗ персональные данные определяются как любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу.

С научной точки зрения персональные данные классифицируются по степени их идентифицируемости. Выделяются идентификаторы, квазиидентифи-

каторы и нечувствительные атрибуты [1, 2]. Идентифицирующие данные позволяют однозначно установить личность субъекта (фамилия, имя, номер телефона). Квазиидентификаторы не обеспечивают однозначной идентификации при изолированном рассмотрении, однако их комбинация может привести к установлению личности (дата рождения, место проживания, род деятельности). Третью категорию составляют данные, не обладающие идентифицирующими свойствами даже при их комбинировании [2–4].

Понятие обезличивания персональных данных закреплено в Федеральном законе № 152-ФЗ и определяется как совокупность действий, в результате которых становится невозможным установить принадлежность данных конкретному субъекту без использования дополнительной информации.

1.2. ПОНЯТИЕ И СТРУКТУРА ИНФОРМАЦИОННЫХ СИСТЕМ ПЕРСОНАЛЬНЫХ ДАННЫХ (ИСПДн)

Информационная система персональных данных (ИСПДн) представляет собой совокупность технических, программных и организационных средств, обеспечивающих обработку персональных данных. Архитектура ИСПДн включает уровень данных, уровень приложений и уровень взаимодействия с пользователями.

Отдельным компонентом являются средства защиты информации, направленные на предотвращение несанкционированного доступа, утечки и модификации данных [5–7].

1.3. НОРМАТИВНО-ПРАВОВОЕ РЕГУЛИРОВАНИЕ ОБРАБОТКИ И ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Обработка персональных данных в информационных системах регулируется совокупностью нормативно-правовых актов, устанавливающих требования к их сбору, хранению, использованию и защите. В Российской Федерации ключевыми документами в данной области являются Федеральный закон № 152-ФЗ, Постановление Правительства РФ № 1119, Постановление Правительства РФ № 1154 и Приказ ФСТЭК России № 21, формирующие правовые и технические основы обработки персональных данных, включая требования к их защите и обезличиванию [8–11].

Постановление Правительства РФ № 1154 конкретизирует требования к процессу обезличивания, устанавливая необходимость обеспечения невозможности или существенного затруднения идентификации субъекта персональных данных. Существенным аспектом является требование учета риска повторной идентификации [10].

1.4. ТРЕБОВАНИЯ К ОБЕЗЛИЧИВАНИЮ ПЕРСОНАЛЬНЫХ ДАННЫХ В ИСПДн

Ключевым требованием является обеспечение баланса между уровнем защиты и сохранением полезности данных. Методы обезличивания должны выбираться таким образом, чтобы минимизировать потери информации при обеспечении требуемого уровня защиты.

Важным требованием является комплексный характер обработки данных. Необходимо учитывать все категории персональных данных, включая квази-идентификаторы, а также взаимосвязи между ними.

Дополнительным требованием является контекстно зависимый выбор методов обезличивания. Универсального подхода к обезличиванию персональных данных не существует, и поэтому выбор методов должен осуществляться с учетом конкретных условий обработки данных.

1.5. АНАЛИЗ BITRIX24 КАК ИНФОРМАЦИОННОЙ СИСТЕМЫ ПЕРСОНАЛЬНЫХ ДАННЫХ

Bitrix24 аккумулирует как идентифицирующие, так и поведенческие данные пользователей. Основными объектами обработки являются контакты, компании, сделки и лиды. Между указанными сущностями формируются логические связи.

Каждая из указанных сущностей может содержать поля, включающие персональные данные. Для сущности «контакт» характерны прямые идентификаторы (фамилия, имя, номер телефона, email), а также квазиидентификаторы (должность, регион, подразделение, дата рождения) [2, 3].

Особенностью CRM-систем является высокая степень связности данных, что существенно увеличивает риск повторной идентификации, поскольку позволяет восстанавливать информацию о субъекте на основе анализа связей и косвенных признаков.

Дополнительным фактором риска является возможность экспорта данных из системы, что повышает вероятность их несанкционированного использования.

2. АНАЛИЗ МЕТОДОВ ОБЕСПЕЧИВАНИЯ И ПОСТАНОВКА ЗАДАЧИ РАЗРАБОТКИ ПО

2.1. ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Одной из базовых моделей обезличивания является k -анонимность, предложенная Л. Суини. В рамках такого подхода набор данных считается

защищенным, если каждая запись неотличима как минимум от $k - 1$ других записей по набору квазиидентификаторов. Для достижения этого свойства используются операции обобщения и подавления значений. Несмотря на широкое распространение, k -анонимность обладает рядом ограничений, связанных с возможностью раскрытия чувствительной информации при недостаточном разнообразии данных [2].

С целью устранения указанных недостатков была предложена модель l -разнообразия, обеспечивающая наличие достаточного количества различных значений чувствительных атрибутов в каждой группе. Такой подход снижает риск атак, связанных с однородностью данных, однако не полностью устраняет возможность утечки информации [12].

Отдельное направление представляют методы дифференциальной приватности, основанные на добавлении случайного шума к результатам обработки данных. Такой подход обеспечивает формальные гарантии защиты, однако его применение ограничено в задачах, требующих сохранения точной структуры данных, например в системах класса CRM [13].

Наряду с формальными моделями, на практике применяются маскирование, псевдонимизация, агрегация и обобщение. Псевдонимизация занимает особое место, поскольку позволяет сохранить связность данных, но не исключает возможность восстановления исходной информации при наличии дополнительных данных.

Анализ показывает, что ни один из методов не обеспечивает универсального решения задачи обезличивания [14, 1].

2.2. АНАЛИЗ ПРИМЕНИМОСТИ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ В ИСПДн BITRIX24

Bitrix24 характеризуется наличием разнородных данных, включающих идентифицирующие атрибуты, квазиидентификаторы и неструктурированные текстовые данные. Данные обладают высокой степенью связности, что существенно усложняет применение классических методов.

Маскирование может применяться для частичного сокрытия отдельных атрибутов, но не обеспечивает надежной защиты при атаках на основе сопоставления данных. Псевдонимизация сохраняет связность данных, но допускает восстановление исходных данных при наличии дополнительной информации. Методы обобщения эффективны для табличных данных, но ограничено применимы в системах со сложными взаимосвязями.

Применение отдельных методов в изолированном виде не позволяет обеспечить требуемый уровень защиты.

2.3. АНАЛИЗ РИСКОВ ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Основным источником риска являются квазиидентификаторы. Классические исследования показывают, что 87 % населения США можно идентифицировать по комбинации почтового индекса, даты рождения и пола [3].

Риск повторной идентификации особенно высок в связанных данных. В CRM-системе один субъект может быть представлен в нескольких сущностях: контакт, сделка, задача, комментарий. Даже при удалении явных идентификаторов сохранение связей между сущностями и датами событий создает возможность для повторной идентификации на основе анализа временных паттернов и структуры взаимодействий.

2.4. ОГРАНИЧЕНИЯ И НЕДОСТАТКИ СУЩЕСТВУЮЩИХ РЕШЕНИЙ ОБЕЗЛИЧИВАНИЯ

Ключевым недостатком является ограниченная применимость существующих подходов. Большинство методов разрабатывались для обработки структурированных табличных данных и не учитывают сложные взаимосвязи между объектами.

Фундаментальной проблемой является компромисс между уровнем защиты и сохранением полезности данных. В существующих решениях этот компромисс часто не учитывается должным образом.

Серьезные ограничения связаны с обработкой неструктурированных данных. Большинство методов ориентированы на структурированные данные и не обеспечивают корректного выявления персональных сведений в текстах.

Отсутствие адаптивности методов также является существенным недостатком. В большинстве случаев применяются статические правила обработки, не учитывающие особенности конкретного набора данных и уровень риска.

2.5. ПОСТАНОВКА ЗАДАЧИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБЕЗЛИЧИВАНИЯ ДАННЫХ В BITRIX24

Целью настоящей работы является разработка программного обеспечения для обезличивания персональных данных в системе Bitrix24.

Формально задача может быть представлена следующим образом. Пусть задан исходный набор данных D , экспортированный из Bitrix24. Требуется построить функцию обезличивания F , в результате применения которой формируется обезличенный набор данных:

$$D' = F(D).$$

При этом преобразование должно удовлетворять следующим условиям:

$$R(D') = R_{\text{доп}},$$

$$U(D') = U_{\text{мин}},$$

$$S(D') = S(D),$$

где $R(D')$ – риск повторной идентификации субъекта персональных данных, $R_{\text{доп}}$ – допустимое пороговое значение риска, $U(D')$ – полезность обезличенных данных, $U_{\text{мин}}$ – минимально допустимый уровень полезности для решения прикладных задач, $S(D')$ – структура обезличенного набора данных и логические связи между сущностями Bitrix24.

Функция преобразования FFF должна учитывать типы данных и их классификацию, взаимосвязи между сущностями, контекст использования данных, а также требования нормативных документов. Необходимость учета квази-идентификаторов и риска повторной идентификации обусловлена тем, что даже после удаления прямых идентификаторов возможна идентификация субъекта на основе комбинации косвенных признаков и внешних источников данных [2–4].

Входными данными для разрабатываемой системы являются наборы персональных данных, полученные из Bitrix24 в структурированном и неструктурированном виде. Выходными данными являются обезличенные наборы данных, в которых риск идентификации субъектов снижен до допустимого уровня при сохранении аналитической и функциональной ценности данных.

К основным ограничениям задачи относятся:

- соответствие требованиям нормативных документов в области защиты персональных данных;
- необходимость обработки разнородных данных;
- сохранение целостности и связности данных;
- обеспечение масштабируемости обработки;
- соблюдение требований безопасности, установленных для ИСПДн [9, 11].

Ключевым элементом решения является необходимость оценки риска повторной идентификации. В соответствии с современными подходами этот риск рассматривается как вероятностная характеристика и должен учитываться при выборе методов обработки данных [14, 1].

3. РАЗРАБОТКА И РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБЕЗЛИЧИВАНИЯ ДАННЫХ

3.1. ТРЕБОВАНИЯ К РАЗРАБАТЫВАЕМОМУ ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ

Программное обеспечение должно обеспечивать снижение риска повторной идентификации субъектов при сохранении полезности данных. Система должна учитывать особенности структуры данных Bitrix24, включая наличие идентифицирующих атрибутов, квазиидентификаторов и неструктурированных текстовых данных.

К основным функциональным требованиям системы относится поддержка процессов получения, обработки и сохранения данных, поступающих из Bitrix24. Программное обеспечение должно обеспечивать подключение к CRM-системе с использованием API Bitrix24, получение CRM-сущностей, извлечение метаданных полей и обработку экспортированных наборов данных. Поддержка интеграции с внешней системой является необходимым условием автоматизации процесса обезличивания и минимизации ручного вмешательства в обработку данных.

Система должна обеспечивать обработку основных сущностей Bitrix24, содержащих персональные данные (контакты, лиды, сделки и компании), и применять к ним набор правил обезличивания для стандартных и пользовательских полей.

Одним из ключевых функциональных требований является поддержка нескольких методов обезличивания персональных данных, таких как маскирование, удаление, псевдоминимизация, хэширование, генерация синтетических значений, обобщение данных.

Система должна обеспечивать автоматизированный запуск ETL-процессов, включающих этапы извлечения данных, их нормализации, преобразования и сохранения результатов. Для выполнения длительных операций обработки необходимо использование отдельного механизма оркестрации задач, позволяющего выполнять обработку асинхронно без блокировки пользовательских запросов.

Одним из обязательных требований является ведение журнала аудита действий пользователей и процессов обработки данных.

К нефункциональным требованиям относится обеспечение безопасности, расширяемости системы и воспроизводимости результатов обработки.

3.2. РАЗРАБОТКА МОДЕЛИ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Исходный набор данных представляется в виде множества записей:

$$D = \{d_1, d_2, \dots, d_n\},$$

где D – исходный набор данных, d_i – отдельная запись, n – количество записей. Каждая запись содержит набор атрибутов:

$$d_i = \{a_1, a_2, \dots, a_m\},$$

где a_j – отдельный атрибут записи, m – количество атрибутов.

В рамках модели множество атрибутов разделяется на четыре категории:

$$A = \{ID, QID, SA, SID\},$$

где ID – прямые идентификаторы, QID – квазиидентификаторы, SA – чувствительные атрибуты, SID – служебные и связующие идентификаторы.

Таблица 1

Table 1

Соответствие полей Bitrix24 категориям данных и методам обезличивания

Mapping of Bitrix24 fields to data categories and anonymization methods

Сущность Bitrix24	Поле	Категория	Метод обезличивания	Результат обработки
Контакт	ФИО	ID	Псевдоминимизация	Замена на условное значение, например «Клиент_001»
Контакт	Телефон	ID	Маскирование	Частичное сокрытие номера
Контакт	Email	ID	Защищенное хеширование или псевдонимизация	Исключение прямого восстановления адреса
Контакт	Дата рождения	QID	Обобщение	Замена точной даты возрастным диапазоном
Контакт	Город / регион	QID	Группировка	Замена точного значения укрупненной категорией

Окончание табл. 1

End of the Tab. 1

Сущность Bitrix24	Поле	Категория	Метод обезличивания	Результат обработки
Пользователь	Должность	<i>QID</i>	Обобщение	Замена конкретной должности группой должностей
Пользователь	Подразделение	<i>QID</i>	Группировка	Замена на более общий организационный уровень
Компания	Название компании	<i>QID / SID</i>	Псевдонимизация или обобщение	Замена на условное название или категорию
Сделка	Сумма сделки	<i>SA</i>	Обобщение	Замена точного значения диапазоном
Сделка	Коммерческие условия	<i>SA</i>	Подавление или маскирование	Соккрытие конфиденциальных фрагментов
Комментарий	Текст комментария	<i>SA</i>	Маскирование или подавление	Удаление или замена персональных фрагментов текста
CRM-сущность	ID контакта, сделки, компании	<i>SID</i>	Согласованная псевдонимизация	Сохранение связей между сущностями

Для прямых идентификаторов используется наиболее строгая обработка. Для квазиидентификаторов применяется снижение детализации (замена даты рождения возрастным диапазоном, города – регионом). Чувствительные атрибуты обрабатываются с учетом их значения для последующего анализа. Служебные идентификаторы не удаляются полностью, для них используется согласованная псевдонимизация.

3.3. РАЗРАБОТКА АЛГОРИТМА АДАПТИВНОГО ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Основная задача алгоритма адаптивного обезличивания персональных данных заключается в автоматизированной классификации полей CRM-сущностей Bitrix24, выборе подходящего метода обработки и формировании

обезличенного набора данных с сохранением логических связей между сущностями.

Классификация полей выполняется на основе анализа имени поля, его типа, метаданных и набора эвристических признаков. Для каждого поля вычисляется коэффициент уверенности классификации:

$$\text{conf}(f_i),$$

где f_i – анализируемое поле, $\text{conf}(f_i)$ – коэффициент уверенности в принадлежности поля к определенной категории данных. Чем больше количество совпадений с признаками определенной категории, тем выше значение данного коэффициента. Если коэффициент уверенности превышает заданное пороговое значение, полю назначается соответствующая категория.

Выбор метода осуществляется на основе категории данных и необходимости сохранения аналитической значимости поля. Преобразование данных выполняется последовательно для каждой записи исходного набора:

$$d'_i = T(d_i, p_j),$$

где d_i – исходная CRM-запись, p_j – выбранное правило обработки, d'_i – преобразованная запись.

После выполнения преобразования алгоритм проводит проверку результата, оценивая остаточный риск повторной идентификации. Если риск превышает допустимый уровень, алгоритм усиливает степень обезличивания.

Алгоритм предусматривает режим предварительного просмотра, позволяющий проверить корректность классификации полей до обработки всего набора данных.

3.4. ПРОЕКТИРОВАНИЕ АРХИТЕКТУРЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Архитектура разрабатываемой системы построена на модульном многоуровневом принципе, обеспечивающем безопасную обработку данных, разделение ответственности и возможность расширения. Выбранный подход позволяет отделить пользовательский интерфейс от прикладной логики, логику – от доступа к данным, а интеграцию с Bitrix24 – от алгоритмов обезличивания.

Т а б л и ц а 2

T a b l e 2

Основные компоненты архитектуры программного обеспечения**Key components of software architecture**

Компонент	Назначение
API-слой	Обработка HTTP-запросов, маршрутизация, валидация входных данных
Слой прикладной логики	Управление профилями обезличивания, правилами обработки, запуском задач и экспортом результатов
Модуль обезличивания	Реализация модели и алгоритма адаптивного обезличивания
Модуль интеграции с Bitrix24	Получение CRM-сущностей, метаданных и пользовательских полей из Bitrix24
ETL-подсистема	Выполнение длительных задач извлечения, преобразования и сохранения данных
Слой доступа к данным	Изоляция операций чтения и записи от прикладной логики
Хранилище данных	Хранение профилей, правил, параметров задач, журналов и обезличенных результатов
Подсистема аутентификации	Управление доступом пользователей к функциям системы
Подсистема аудита	Фиксация действий пользователей и этапов обработки без сохранения исходных персональных данных

С точки зрения безопасности архитектура предусматривает следующие меры:

- разграничение доступа пользователей по ролям;
- изоляцию модуля интеграции с Bitrix24 от остальных компонентов;
- отсутствие исходных персональных данных в системных логах;
- хранение параметров подключений в защищенном виде;
- журналирование операций обработки;
- выполнение длительных задач в отдельной ETL-подсистеме;
- сохранение только обезличенных результатов для дальнейшего анализа.

Выбранная архитектура обеспечивает модульность программного обеспечения, независимость компонентов, возможность масштабирования обработки данных и расширения набора методов обезличивания. Кроме того, она позволяет разделить обработку пользовательских запросов и длительные ETL-задачи, что повышает устойчивость системы при работе с большими объемами CRM-данных.

3.5. РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Реализация программного обеспечения выполнена на языке Python с использованием асинхронного подхода и модульной структуры проекта. Серверная часть реализована на основе FastAPI [15]. Для описания структур данных применяются схемы Pydantic.

Модуль обезличивания реализует операции маскирования, псевдонимизации, обобщения, подавления и защищенного преобразования. Реализован модуль сопоставления полей для автоматического анализа стандартных и пользовательских атрибутов.

Интеграция с Bitrix24 реализована с использованием библиотеки httpx для асинхронных HTTP-запросов; ETL-задачи выполняются в Apache Airflow. Для разграничения доступа используется проверка JWT-токенов от Keycloak.

4. ОЦЕНКА ЭФФЕКТИВНОСТИ И РЕКОМЕНДАЦИИ

4.1. ПОДГОТОВКА ТЕКСТОВЫХ ДАННЫХ

В ходе разработки программного обеспечения был сформирован синтетический тестовый набор данных, моделирующий структуру CRM-данных Bitrix24. Набор включает основные CRM-сущности: контакты, лиды, сделки, компании – по 100 записей для каждой сущности. Данные содержат прямые идентификаторы, квазиидентификаторы, пользовательские поля и неструктурированные текстовые комментарии.

4.2. МЕТОДИКА ТЕСТИРОВАНИЯ И КРИТЕРИИ ОЦЕНКИ ЭФФЕКТИВНОСТИ

Оценка эффективности разрабатываемого программного обеспечения проводилась на основе сравнения двух состояний набора данных: исходного

и после обезличивания. Основным показателем являлось изменение риска повторной идентификации:

$$\Delta R = R(D) - R(D'),$$

где $R(D)$ – риск повторной идентификации в исходном наборе данных, $R(D')$ – риск повторной идентификации после обезличивания, ΔR – величина снижения риска.

Риск оценивается на основе анализа уникальности записей по квазиидентификаторам. Риск отдельной записи определяется следующим образом:

$$R_i = \frac{1}{|G_i|},$$

где R_i – риск повторной идентификации отдельной записи, G_i – группа записей с одинаковыми значениями квазиидентификаторов, $|G_i|$ – количество записей в данной группе.

Общий риск для набора данных рассчитывался как доля записей, отнесенных к группе повышенного риска:

$$R(G) = \frac{N_{risk}}{N_{total}} 100 \%,$$

где N_{risk} – количество записей с высоким риском повторной идентификации, N_{total} – общее количество записей в наборе данных.

После применения алгоритма аналогичный расчет выполнялся для обезличенного набора данных D' . Сравнение значений $R(D)$ и $R(D')$ позволяло определить степень снижения риска повторной идентификации.

4.3. РЕЗУЛЬТАТЫ РАБОТЫ АЛГОРИТМА ОБЕЗЛИЧИВАНИЯ

В результате применения разработанного алгоритма адаптивного обезличивания был сформирован обезличенный набор CRM-данных, включающий сведения по четырем основным сущностям Bitrix24: контактам, лидам, сделкам и компаниям. Обработка выполнялась на синтетическом тестовом наборе данных, подготовленном в разделе 4.1.

В ходе работы алгоритма было обработано 400 записей. Результаты обработки представлены в табл. 3. Стоит отметить, что применение согласованной псевдонимизации позволило сохранить все логические связи между CRM-сущностями.

Таблица 3

Table 3

Сводные результаты обработки текстовых данных**Summary results of text data processing**

CRM-сущность	Количество входных записей	Количество обработанных записей	Количество полей в сущности	Количество полей, подлежащих обезличиванию	Обработано полей	Ошибки
Контакты	100	100	36	19	19	0
Лиды	100	100	34	18	18	0
Сделки	100	100	32	14	14	0
Компании	100	100	30	12	12	0
Итого	400	400	132	63	63	0

4.4. ОЦЕНКА РИСКА ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ПОСЛЕ ОБЕЗЛИЧИВАНИЯ

После завершения работы алгоритма обезличивания была проведена оценка риска повторной идентификации данных в соответствии с методикой, изложенной в разделе 4.2. Результаты оценки риска представлены в табл. 4.

Таблица 4

Table 4

Сравнение риска повторной идентификации до и после обезличивания**Comparison of re-identification risk before and after anonymization**

CRM-сущность	Количество записей	Риск до обработки $R(D)$, %	Риск после обработки $R(D')$, %	Снижение риска ΔR , %
Контакты	100	82	24	58
Лиды	100	78	26	52
Сделки	100	64	22	42
Компании	100	56	18	38
Итого	400	70	22,5	47,5

Наибольшее снижение риска наблюдается для сущностей, содержащих большое количество прямых идентификаторов и квазиидентификаторов.

Остаточный риск полностью не устраняется. Это связано с сохранением части структуры CRM-данных, возможной уникальностью отдельных комбинаций атрибутов и потенциальным использованием внешних источников информации.

4.5. ВЫВОДЫ ПО ЭФФЕКТИВНОСТИ РАЗРАБАТЫВАЕМОГО РЕШЕНИЯ И РЕКОМЕНДАЦИИ ПО ПРИМЕНЕНИЮ

Разработанное программное обеспечение выполняет автоматизированное обезличивание CRM-данных Bitrix24, снижает риск повторной идентификации и сохраняет структуру данных.

Практическая ценность заключается в возможности централизованного управления процессом обезличивания. Пользователь может настраивать профили обработки, изменять правила, выполнять предварительный просмотр и запускать ETL-процессы.

При применении разработанного ПО рекомендуется использовать отдельные профили для разных сценариев, выполнять предварительный просмотр, проверять пользовательские поля, ограничивать доступ к исходным данным и журналам, регулярно обновлять словари признаков ПДн.

Для дальнейшего развития целесообразно внедрение количественных метрик оценки риска, реализация проверки k -анонимности, расширение эвристик для пользовательских полей, развитие обработки неструктурированных текстовых данных, внедрение административного web-интерфейса.

ЗАКЛЮЧЕНИЕ

В статье представлен прототип программного обеспечения для адаптивного обезличивания персональных данных в системе Bitrix24, позволяющий снизить риск повторной идентификации при сохранении структуры и прикладной ценности CRM-данных. Разработанное решение поддерживает настройку профилей и правил обработки, классификацию стандартных и пользовательских полей, применение методов маскирования, псевдонимизации, обобщения, подавления и защищённого преобразования.

Результаты экспериментальной оценки подтверждают эффективность предложенного подхода: риск повторной идентификации снижен с 70 до 22,5 %, при этом алгоритм корректно обрабатывает различные категории персональных данных.

К числу ограничений прототипа относится отсутствие полной реализации формальных моделей k -анонимности, l -разнообразия и t -близости, а также зависимость качества обработки неструктурированных текстовых данных от полноты правил.

Перспективы дальнейшего развития включают внедрение формализованной проверки k -анонимности, расширение метрик оценки риска, совершенствование обработки текстовых данных и разработку административного веб-интерфейса.

СПИСОК ЛИТЕРАТУРЫ

1. Privacy-preserving data publishing: A survey of recent developments / B.C.M. Fung, K. Wang, R. Chen, P.S. Yu // *ACM Computing Surveys*. – 2010. – Vol. 42 (4). – P. 1–53.
2. Sweeney L. k -Anonymity: A model for protecting privacy // *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. – 2002. – Vol. 10 (5). – P. 557–570.
3. Sweeney L. Simple demographics often identify people uniquely. – Pittsburgh, 2000. – URL: <https://dataprivacylab.org/projects/identifiability/> (accessed: 08.09.2026).
4. Narayanan A., Shmatikov V. Robust De-anonymization of Large Sparse Datasets // *IEEE Symposium on Security and Privacy*. – IEEE, 2008. – P. 111–125.
5. Назаров Д.М., Саматов К.М. Основы обеспечения безопасности персональных данных. – Екатеринбург: УрФУ, 2019. – 216 с.
6. Обеспечение защиты персональных данных: методическое пособие / И.А. Баймакова, А.В. Новиков, А.И. Рогачев, А.Х. Хыдыров. – 3-е изд. – М.: 1С-Паблишинг, 2011. – 268 с.
7. Шаньгин В.Ф. Информационная безопасность компьютерных систем и сетей: учебное пособие. – М.: Форум, 2021. – 416 с.
8. О персональных данных: Федеральный закон РФ от 27.07.2006 № 152-ФЗ // *Российская газета*. – 2006. – 29 июля.
9. Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных: постановление Правительства РФ от 01.11.2012 № 1119 // *Собрание законодательства Российской Федерации*. – 2012. – № 45. – Ст. 6257.
10. Об утверждении требований к обезличиванию персональных данных, методов обезличивания персональных данных и Правил обезличивания персональных данных: постановление Правительства РФ от 01.08.2025 № 1154 // *Российская газета*. – 2025. – № 170. – С. 3.
11. Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных: приказ ФСТЭК России от 18.02.2013 № 21 // *Российская газета*. – 2013. – 17 мая.

12. L-diversity: Privacy beyond k-anonymity / A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam // ACM Transactions on Knowledge Discovery from Data. – 2007. – Vol. 1 (1). – P. 1–52.
13. Dwork C. Differential privacy // Automata, Languages and Programming. ICALP 2006. – Berlin: Springer, 2006. – P. 1–12. – (Lecture Notes in Computer Science; vol. 4052).
14. El Emam K. Guide to the de-Identification of personal health information. – Boca Raton: CRC Press, 2013. – 384 p.
15. FastAPI Documentation. – URL: <https://fastapi.tiangolo.com/> (accessed: 08.09.2026).

Синюшкин Тимофей Андреевич, лаборант кафедры защиты информации Новосибирского государственного технического университета. E-mail: kaf_zi@corp.nstu.ru

Исаев Глеб Андреевич, старший преподаватель кафедры защиты информации Новосибирского государственного технического университета. Основное направление научных исследований – булевы функции в теории кодирования и криптографии. E-mail: g.isaev@corp.nstu.ru

DOI: 10.17212/2782-2230-2026-3-64-83

Development of software for personal data anonymization in the Bitrix24 personal data information system*

T.A. Sinyushkin¹, G.A. Isaev²

¹ *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, laboratory assistant at the Department of Information Security. E-mail: kaf_zi@corp.nstu.ru*

² *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, Lecturer of the Department of Information Security. E-mail: g.isaev@corp.nstu.ru*

This article discusses the development of software for anonymizing personal data within the Bitrix24 personal data information system. The study is motivated by the need to utilize CRM system data for testing, analytics, and training purposes while simultaneously complying with the requirements of Federal Law No. 152-FZ and Russian Government Decree No. 1154. An adaptive approach is proposed, based on classifying CRM entity attributes into

* Received 16 July 2026.

four categories: direct identifiers, quasi-identifiers, sensitive attributes, and service attributes. An algorithm has been developed for the dynamic selection of anonymization methods (masking, pseudonymization, generalization, and suppression) based on field category and estimated re-identification risk. A modular software architecture is presented, comprising a Bitrix24 integration module, an anonymization module, and an ETL subsystem. Experimental evaluation using synthetic data demonstrated a reduction in re-identification risk from 70 % to 22.5 % while preserving 100 % of the logical relationships between entities.

Keywords: personal data, anonymization, Bitrix24, CRM, personal data information system, quasi-identifiers, re-identification, pseudonymization, masking, ETL, API, information protection

REFERENCES

1. Fung B.C.M., Wang K., Chen R., Yu P.S. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys*, 2010, vol. 42 (4), pp. 1–53.
2. Sweeney L. k-Anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002, vol. 10 (5), pp. 557–570.
3. Sweeney L. *Simple demographics often identify people uniquely*. Pittsburgh, 2000. Available at: <https://dataprivacylab.org/projects/identifiability/> (accessed 08.09.2026).
4. Narayanan A., Shmatikov V. Robust De-anonymization of Large Sparse Datasets. *IEEE Symposium on Security and Privacy*, 2008, pp. 111–125.
5. Nazarov D.M., Samatov K.M. *Osnovy obespecheniya bezopasnosti personal'nykh dannykh* [Fundamentals of personal data security]. Yekaterinburg, UrFU Publ., 2019. 216 p.
6. Baimakova I.A., Novikov A.V., Rogachev A.I., Khydyrov A.Kh. *Obespechenie zashchity personal'nykh dannykh* [Ensuring personal data protection]. 3rd ed. Moscow, 1C-Publishing, 2011. 268 p.
7. Shan'gin V.F. *Informatsionnaya bezopasnost' komp'yuternykh sistem i setei* [Information security of computer systems and networks]. Moscow, Forum Publ., 2021. 416 p.
8. On personal data: Federal Law of the Russian Federation No. 152-FZ dated July 27, 2006. *Rossiyskaya gazeta = Russia's Newspaper*, 2006, July 29. (In Russian).
9. On approving requirements for personal data protection during processing in personal data information systems: Decree of the Government of the Russian Federation No. 1119 dated November 1, 2012. *Sobranie zakonodatel'stva Rossiiskoi Federatsii = Collection of the legislation of the Russian Federation*, 2012, no. 45, Art. 6257. (In Russian).

10. On approval of requirements for the de-identification of personal data, methods for the de-identification of personal data, and Rules for the de-identification of personal data: Decree of the Government of the Russian Federation No. 1154 dated August 1, 2025. *Rossiyskaya gazeta = Russia's Newspaper*, 2025, no. 170, p. 3. (In Russian).
11. On approval of the composition and content of organizational and technical measures to ensure the security of personal data during processing in personal data information systems: Order of the FSTEC of Russia No. 21 dated February 18, 2013. *Rossiyskaya gazeta = Russia's Newspaper*, 2013, May 17. (In Russian).
12. Machanavajjhala A., Kifer D., Gehrke J., Venkatasubramanian M. L-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data*, 2007, vol. 1 (1), pp. 1–52.
13. Dwork C. Differential privacy. *Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science*, vol. 4052. Berlin, Springer, 2006, pp. 1–12.
14. El Emam K. *Guide to the de-identification of personal health information*. Boca Raton, CRC Press, 2013. 384 p.
15. *FastAPI documentation*. Available at: <https://fastapi.tiangolo.com/> (accessed 08.09.2026).

Для цитирования:

Синюшкин Т.А., Исаев Г.А. Разработка программного обеспечения для обезличивания персональных данных в ИСПДн Bitrix24 // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 64–83. – DOI: 10.17212/2782-2230-2026-3-64-83.

For citation:

Sinyushkin T.A., Isaev G.A. Razrabotka programmnoy obespecheniya dlya obezlichivaniya personal'nykh dannykh v ISPDn Bitrix24 [Development of software for personal data anonymization in the Bitrix24 personal data information system]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 64–83. DOI: 10.17212/2782-2230-2026-3-64-83.