

# Извлечение и обработка данных из систем управления данными пациентов в отделениях интенсивной терапии

**Источник:** Frontiers in Digital Health

**Оригинал:** <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1747227>

ETL

GDPR

Python

интенсивная терапия

информационные системы

обработка данных

регулирование

## Введение

Рутинные клинические данные, фиксируемые в системах управления данными пациентов (**PDMS** — Patient Data Management Systems) в условиях интенсивной терапии и периоперационного периода, являются бесценным ресурсом для клинических исследований. Однако проприетарная, фрагментированная и ориентированная на транзакции архитектура многих систем серьезно ограничивает вторичное использование данных и требует обширной обработки в рамках процессов **ETL** (Extract, Transform, and Load — извлечение, преобразование и загрузка).

## Методы

Мы разработали модульный ETL-фреймворк на базе языка **Python**, который обеспечивает гибкое, специфичное для конкретной области извлечение высокочастотных мультимодальных данных из **PDMS**. Система предоставляет многоуровневые компоненты для извлечения, предобработки, гармонизации и деидентификации данных, что позволяет адаптировать или расширять методы извлечения без изменения базовой архитектуры. Каждая клиническая область представлена с помощью специализированных моделей

**Pydantic**, обеспечивающих согласованные схемы выходных данных, ограничения типов и автоматизированные проверки правдоподобности. **SQLAlchemy** абстрагирует доступ к базам данных, в то время как структурированная логика предобработки устраняет распространенные несоответствия в документации и преобразует гетерогенные записи **PDMS** в стандартизированные представления.

## Результаты

Фреймворк создает воспроизводимые, готовые к анализу наборы данных посредством прозрачного и поддающегося аудиту рабочего процесса. Интегрированный журнал аудита фиксирует параметры извлечения, преобразования и производные поля, обеспечивая полную прослеживаемость. Необратимая псевдонимизация с использованием «соли» (salted pseudonymization) встроена непосредственно в конвейер обработки, что обеспечивает соответствие требованиям Общего регламента по защите данных Европейского союза (**GDPR**; нем. Datenschutz-Grundverordnung, **DSGVO**) и статьи 27 Закона Баварии о больницах (**BayKrG** — Bayerisches Krankenhausgesetz). Инкапсулируя логику извлечения в модульные блоки обработки с согласованной валидацией и автоматизированной деидентификацией, система заменяет сложные разовые запросы стандартизированными, поддерживаемыми и готовыми к исследованиям процессами.

## Заключение

Представленный фреймворк преодолевает существенные технические и нормативные барьеры для вторичного использования данных **PDMS**, внедряя конвейер извлечения, ориентированный в первую очередь на управление данными (governance-first). Его модульная архитектура инкапсулирует специфичные для конкретного учреждения запросы к **PDMS** в ограниченном слое адаптеров, сохраняя при этом валидацию, псевдонимизацию и аудит переносимыми и многократно используемыми в различных областях и установках. Благодаря внедрению моделей валидации на уровне предметных областей, необратимой псевдонимизации и структурированного аудита, фреймворк обеспечивает воспроизводимый, соответствующий нормативным требованиям доступ к высокочастотным данным интенсивной терапии. Вместо требования немедленного приведения к единой модели данных, он предоставляет прагматичную основу, на которой семантическая и синтаксическая интероперабельность могут добавляться постепенно по мере развития требований и ресурсов.

---

---

Перевод выполнен: 10.06.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.