

УДК 004.06

## Разработка сервиса для конвертации и загрузки в базу данных файлов инвентаризации Мухин А.А. (ИТМО)

Научный руководитель — кандидат технических наук, доцент Федоров Д.А. (ИТМО)

**Введение.** В современном мире автоматизация обработки данных играет ключевую роль в повышении эффективности бизнес—процессов. Одним из вызовов стала необходимость автоматизации обработки инвентаризационных данных для Системы Экологического Мониторинга (СЭМ), поступающих в различных форматах. Ручной ввод данных приводит к высокой нагрузке на сотрудников, увеличению вероятности ошибок и низкой масштабируемости системы. Существующие решения, не подходят для работы с проприетарными форматами, что требует разработки специализированного сервиса. Целью работы является создание автоматизированного сервиса для конвертации и загрузки файлов инвентаризации в базу данных СЭМ.

**Основная часть.** Для реализации проекта были решены следующие задачи

1. Описание особенностей структуры файлов инвентаризации:
  - проведен анализ входных данных: JSON, XLSX и проприетарные формат,
  - проведен анализ структуры файлов,
  - определены основные параметры, подлежащие стандартизации.
2. Аналитический обзор структурированных и проприетарных форматов:
  - исследованы достоинства и недостатки различных форматов,
  - изучены особенности файлов различных форматов,
  - разработаны алгоритмы обработки данных, включая парсинг и нормализацию.
3. Разработка алгоритма конвертации файлов в стандартизированный JSON:
  - создан механизм извлечения данных и их трансформации в единый формат,
  - применены методы валидации и нормализации данных.
4. Валидация качества конвертирования (CI/CD, DevOps):
  - разработана система автоматического тестирования корректности преобразования,
  - внедрены механизмы валидации данных на соответствие структуре базы данных, включая проверку обязательных полей, типов данных и диапазонов значений,
  - внедрены инструменты CI/CD для развертывания и тестирования сервиса.
5. Подготовка данных и загрузка в базу данных:
  - разработана функция загрузки данных в базу данных с использованием ORM SQLAlchemy [1],
  - обеспечена поддержка транзакций для гарантии целостности данных.
6. Разработка минимально—жизнеспособного продукта (MVP):
  - создан микросервис на основе FastAPI[2] для конвертации и загрузки файлов в базу данных MySQL[3],
  - разработан веб—интерфейс на Angular[4] для загрузки файлов пользователями,
  - настроена система контейнеризации с использованием Docker[5] и Nginx[6].
7. Разработка схемы модернизации системы:
  - создана схема по улучшению системы,
  - разработана методика работы с новыми типами файлов.

**Выводы.** Разработанный сервис автоматизирует процесс обработки данных инвентаризации, что значительно снижает нагрузку на сотрудников и минимизирует ошибки, вызванные человеческим фактором - например, опечатки [7]. Данное решение повышает точность и скорость загрузки данных в БД. Сервис обеспечивает бесшовную интеграцию с существующей системой экологического мониторинга, повышая эффективность анализа данных и принятия решений. Практическое использование сервиса позволяет ускорить обработку данных и адаптироваться к новым форматам файлов в будущем. Данное решение

уникально и настроено на конкретную специфику предприятия, поскольку решает узкоспециализированную задачу по конвертации специфичных файлов проприетарных форматов.

**Список использованных источников:**

1. Официальная документация SQLAlchemy – [Электронный ресурс] URL: <https://docs.sqlalchemy.org/> (дата обращения: 13.02.2025)
2. Официальная документация FastAPI – [Электронный ресурс] URL: <https://fastapi.tiangolo.com/> (дата обращения: 13.02.2025)
3. Официальная документация Angular – [Электронный ресурс] URL: <https://angular.io/> (дата обращения: 13.02.2025)
4. Официальная документация MySQL – [Электронный ресурс] URL: <https://dev.mysql.com/doc/> (дата обращения: 13.02.2025)
5. Официальная документация Docker – [Электронный ресурс] URL: <https://docs.docker.com/> (дата обращения: 13.02.2025)
6. Официальная документация Nginx – [Электронный ресурс] URL: <https://nginx.org/ru/> (дата обращения: 13.02.2025)
7. Методологический подход к определению влияния человеческого фактора на работоспособность информационных систем – [Электронный ресурс] URL: <https://habr.com/ru/companies/cognitive/articles/209266/> (дата обращения: 13.02.2025)