



Модус: Управление корпоративным хранилищем данных (ETL)

ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК

Программа для ЭВМ «Модус:Управление корпоративным хранилищем данных (ETL)» (далее – «Программный продукт») предназначена для:

- автоматизированного сбора данных из различных источников;
- трансформации данных и управления качеством данных;
- построения корпоративного хранилища данных.

Программный продукт разработан с использованием «1С:Библиотека стандартных подсистем 8» (далее – БСП).

Схема взаимодействия Программного продукта с источниками данных и хранилищем данных:



Схема взаимодействия Программного продукта с источниками и хранилищем данных

## Список функций Программного продукта

### Получение данных из следующих источников:

- учетных систем: 1С, 1CFresh;
- баз данных: MS SQL, PostgreSQL, Oracle;
- веб-сервисов;
- файлов Excel;
- текстовых файлов с разделителями.

### Сбор данных

Многопоточная загрузка данных и управление очередями для обеспечения производительной и автоматизированной (с возможностью обработки ошибок) загрузки.

Для учетных систем на платформе 1С есть возможность получать данные, используя:

- запросы на языке 1С;
- подключение и управление схемой компоновки данных (СКД) 1С-отчета.

Использовать 1С-скрипты для пост-обработки получаемых данных.

Возможность загрузки метаданных о структуре данных в источниках.

Полученные метаданные могут использоваться при написании запросов с использованием конструктора запросов или вручную.

### Трансформация данных

Трансформация данных предполагает объединение, очистку, дополнение, замену (в т.ч. по правилам соответствия/мэппинга), группировку, форматирование и т.п. операции с данными.

Возможность при трансформации данных использовать:

- шаблоны типовых операций трансформации данных (выбора, группировки, мэппинга и пр.), настраиваемые в визуальном конструкторе;
- SQL-скрипты;
- обработки на языке 1С.

Многошаговые сценарии трансформации.

Диаграмма потока (flow chart) для визуализации состава ETL-пакета с возможностью настраивать правила получения и трансформации данных для компонентов ETL-пакета.

## **Автоматизация ETL-процессов**

- Запуск ETL процессов по расписанию;
- Оповещение о статусе выполнения на электронную почту по списку рассылки;
- Логирование ETL процессов;
- Интерфейсы для контроля выполнения ETL-процессов.

## **Управление хранилищем (DWHM)**

Используемые системы управления базами данных:

- MS SQL;
- PostgreSQL.

Операции по управлению хранилищем:

- Трансформация первичных данных в схему "звезда" с генерацией структуры объектов;
- Визуализация структуры хранилища: ER-диаграммы моделей данных на основе загруженных метаданных хранилища;
- Создание витрин данных в визуальном интерфейсе;
- Версионирование данных.

Управление качеством данных (DQM):

- Настройка правил проверки и очистки данных;
- Шаблоны правил проверки и очистки для "доменов данных" (ИНН, телефонов, СНИЛС, адресов и т.д.);
- Настройка правил сопоставления (мэппинга и замены) данных.