**Министерство науки высшего образования Российской Федерации**

**федеральное государственное автономное образовательное учреждение высшего образования**

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»**

**(Университет ИТМО)**

Факультет информационных технологий и программирования

Образовательная программа Программирование и интернет-технологии

Направление подготовки (специальность) 09.03.02 - Информационные системы и технологии

О Т Ч Е Т

об учебной практике

Тема задания: Разработка ETL процесса

Обучающийся: Шарапков Егор Викторович, М33081

Руководитель практики от профильной организации: Шараева Кристина Витальевна, Университет ИТМО, преподаватель практики.

Руководитель практики от университета: Зубок Д.А., доцент ФИТИП, Университет ИТМО

Практика пройдена с оценкой \_\_\_\_

Дата 17.06.2022

Санкт-Петербург

2022

Этап 1

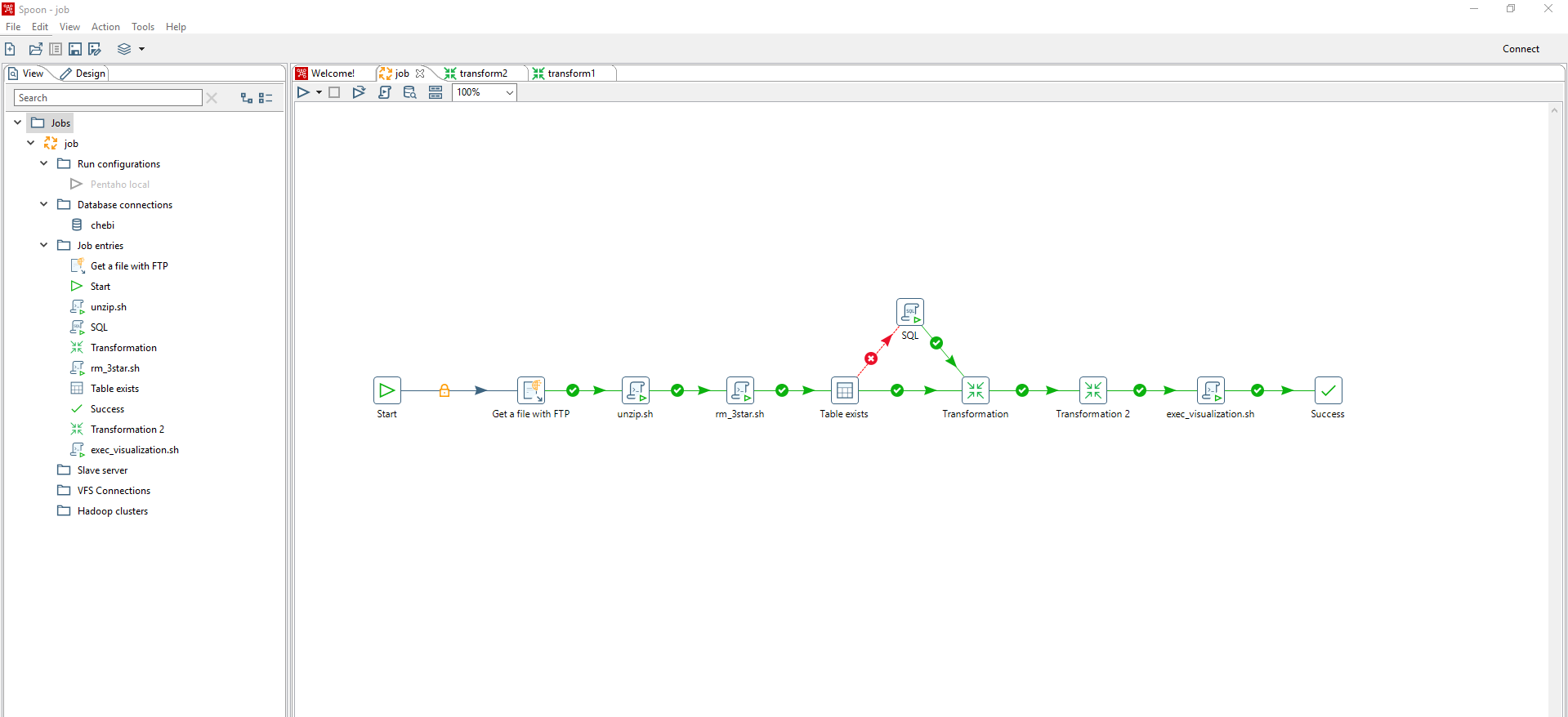
Ознакомление c программным обеспечением Pentaho.

В результате изучения материалов, а именно – прохождения туториалов (руководств), прочтения статей и обсуждений на форумах были получены навыки работы с программным обеспечением Pentaho Data Integration Kettle ETL.

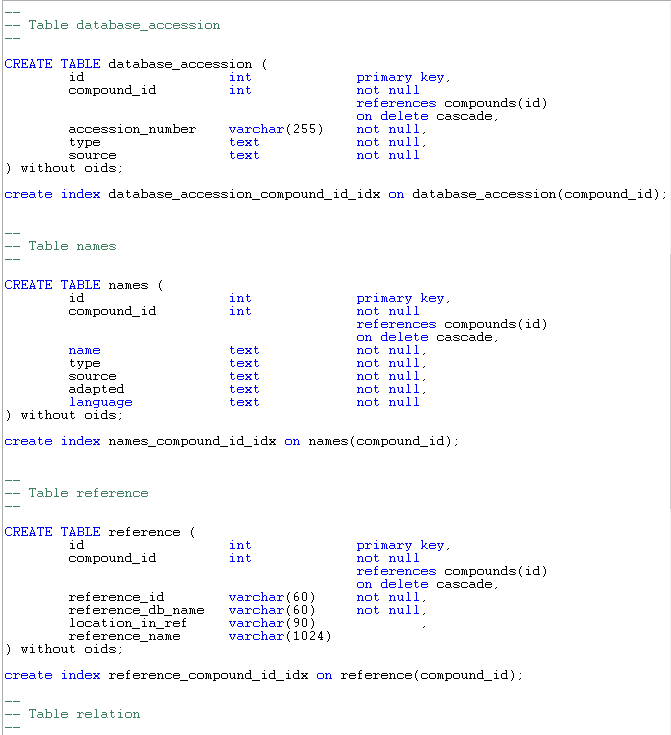
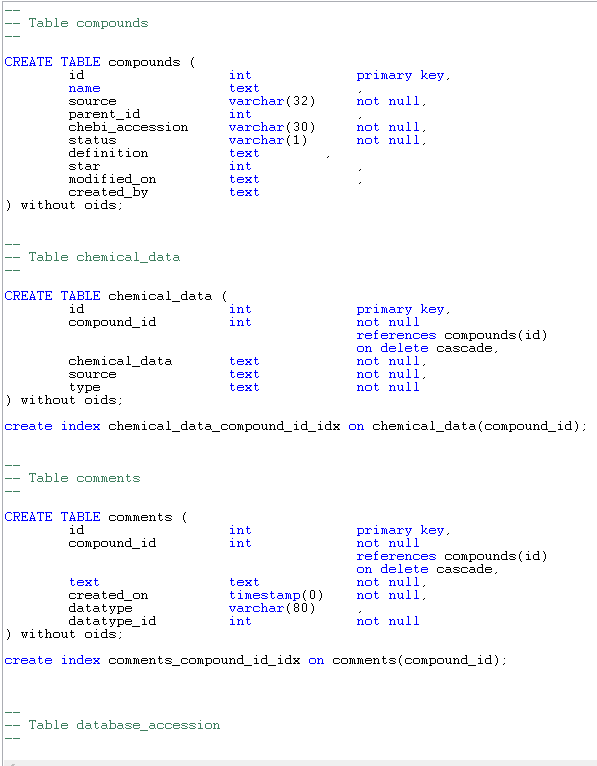
Этап 2

Создание job и схемы БД.

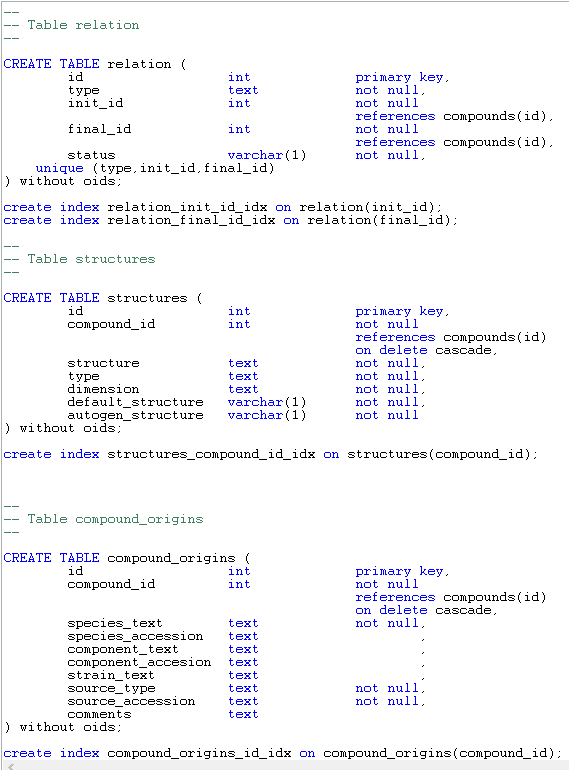
Ссылка на удаленный репозиторий с программным кодом: <https://github.com/EgorGrib/ETL-Process>

Скриншот созданной job в программном обеспечении Pentaho Data Integration Kettle ETL.

SQL скрипт создания схемы БД Postgres.

 2 из 3.

1 из 3.



3 из 3.

Этап 3

Создание трансформаций для парсинга всех данных из файлов и записи их в таблицы.

Трансформация для парсинга файла compounds.tsv, представляющий собой описание химических соединений, и записи данных в таблицу базы данных.

Изображение выглядит как стол

Автоматически созданное описание