# Миграция с SQL Server на PostgreSQL

Миграция может выполняться для целей тестирования, в этом случае необходимо выполнить обезличивание данных.

Обезличивание данных состоит из двух этапов:

1. Анализ данных на наличие персональных данных. Анализ заключается в поиске имен людей и согласованной их замене на обезличенные идентификаторы.
2. Экспорт данных из SQL Server c одновременной заменой и сокрытием персональных данных в скалярных полях и полях, содержащих XML

Экспорт данных может осуществляться без процедуры обезличивания данных.

Экспорт данных может производиться непосредственно из таблиц и из запросов.

При импорте данных используются BULK INSERT (SQL Server) или BINARY COPY (PostgreSQL) для ускорения загрузки данных.

При операции экспорта в БД назначения рекомендуется удалить внешние ключи и все индексы (кроме первичных ключей)

Для целей миграции созданы две утилиты:

1. Finder – поиск имен и присвоение уникальных идентификаторов
2. SqlServerToPostgres – утилита экспорта данных из SQL Server.

Утилиты можно скомпилировать и выполнять как на Linux, так и на Windows.

## Обезличивание и миграция данных

### Finder

Настройки утилиты содержатся в файле appsettings.json

A screen shot of a computer program

Description automatically generated

В настройка необходимо указать строку подключения к БД, шаблон полного имени (SFM|FMS), а также другие технические параметры необходимые для успешного выполнения утилиты.

В результате работы утилита создает файл substitutions.json, который содержит пары имени и соответствующего идентификатора, на который будет производится замена соответствующего имени при экспорте.

A screenshot of a computer program

Description automatically generated

### SqlServerToPostgres

Данная утилита может работать в двух режимах:

1. GenerateExportOptions – в этом режиме автоматически создается файл мэппинга таблиц и полей из БД источника с БД приёмника. Результатом работы утилиты является файл exports.json, данные из которого могут быть использованы в качестве настроек секции Expors (appsettings.json)
2. Export – режим экспорта данных в соответствии с настройками

Для упрощения процедуры экспорта в БД приёмнике рекомендуется удалить все внешние ключи и пересоздать их после завершения процесса миграции данных.

Для обеспечения гибкости и снижения требований к объему БД миграцию данных можно производить частями.

Настройки:

A screen shot of a computer

Description automatically generated

| **Свойство** | **Тип** | **Комментарий** |
| --- | --- | --- |
| **OperationsType** | строка | GenerateExportOptions | Export |
| **SourceConnectionString** | строка подключения к БД источника данных |  |
| **DestinationConnectionString** | строка подключения к БД приёмника |  |
| **Anonymizer** |  |  |
| **SubstitutionsFileName** | cторока | путь к файлу подстановок (Finder) |
| **NamesXPath** | строка | XPath к элементам потенциально содержащему имена сотрудников. Найденные имена в этих элементах будут заменены на идентификаторы из файла (substitutions.json) |
| **MaskXPath** | строка | XPath к элементам потенциально содержащему конфиденциальную информацию (login, email, password и т.п.). текст в таких элементах будет заменен на «\*» |
| **Exports** | массив объектов |  |
| **DestinationTable** | строка | имя таблицы приёмника данных |
| **SourceTable** | строка | имя таблицы источника данных. |
| **SourceQuery** | строка | SQL запрос – источник данных, если не указана таблица источник |
| **SourceQueryKey** | массив строк | ключи для сортировки и пакетирования данных из запроса |
| **BatchSize** | число | размер пакета |
| **Columns** | массив объектов | мэппинг полей источника и приёмника |
| **Destination** | строка | Поле в таблице приёмнике |
| **Source** | строка | Поле в таблице|запросе источнике |
| **Action** | перечисление | Skip – пропускает поле,  Copy – копирует значение из источника в приёмник,  Convert – использует специальный конвертер (реализовано XmlAnonymizerService и TextAnonymizerService)  TypeConvert – производит конвертацию типа данных,  SetNull – устанавливает значение NULL,  MaskWithChar – маскирует поле «\*» |
| **Parameter** | строка | Параметр зависит от Action |
|  |  |  |

Пример результата работы SqlServerToPostgress

A computer screen with green text

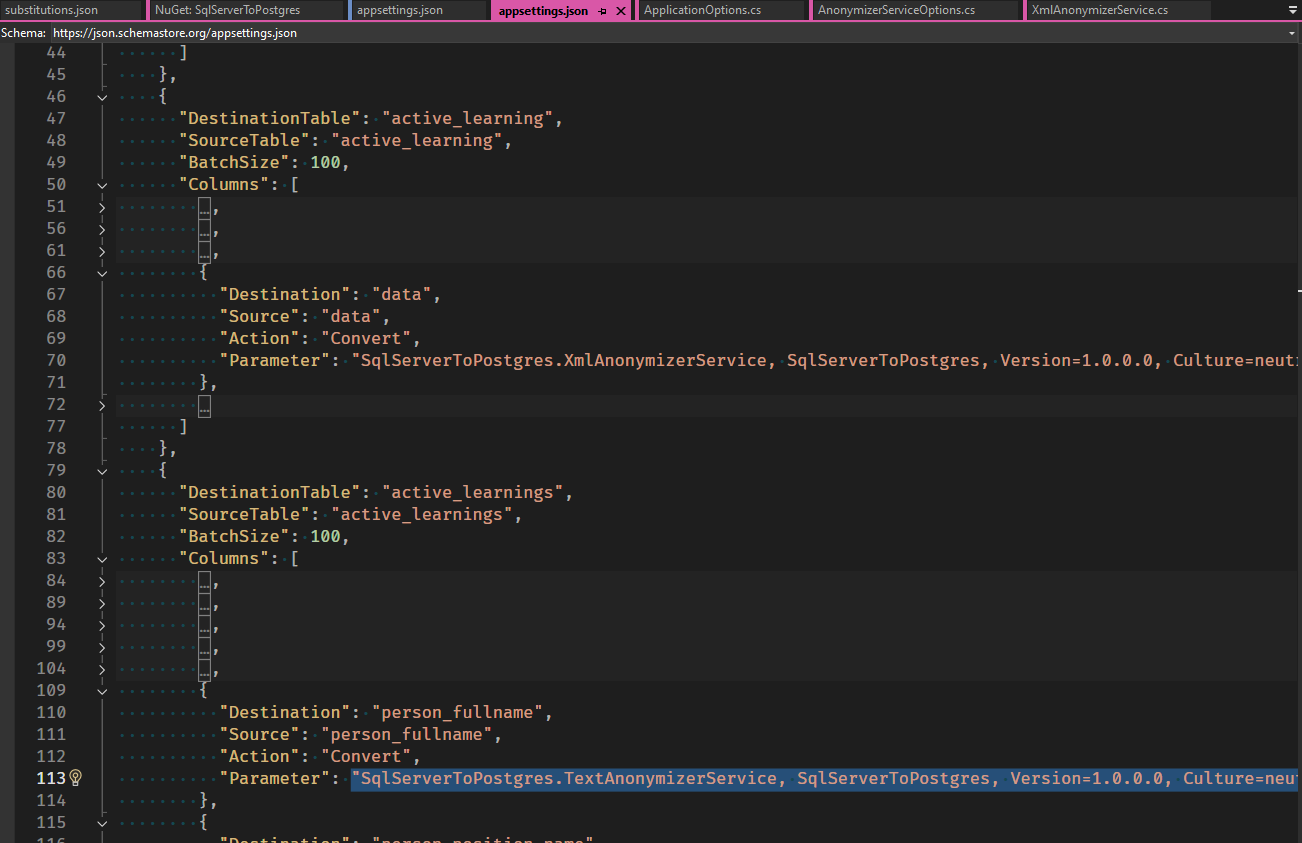
Description automatically generated

Для обезличивания полей, содержащих XML, используется конвертер:

"SqlServerToPostgres.XmlAnonymizerService, SqlServerToPostgres, Version=1.0.0.0, Culture=neutral, PublicKeyToken=null"

Для обезличивания скалярных полей используется конвертер:

"SqlServerToPostgres.TextAnonymizerService, SqlServerToPostgres, Version=1.0.0.0, Culture=neutral, PublicKeyToken=null"



Если обезличивание не требует, в настройках Action = “Convert” следует заменить на Action = “Copy”