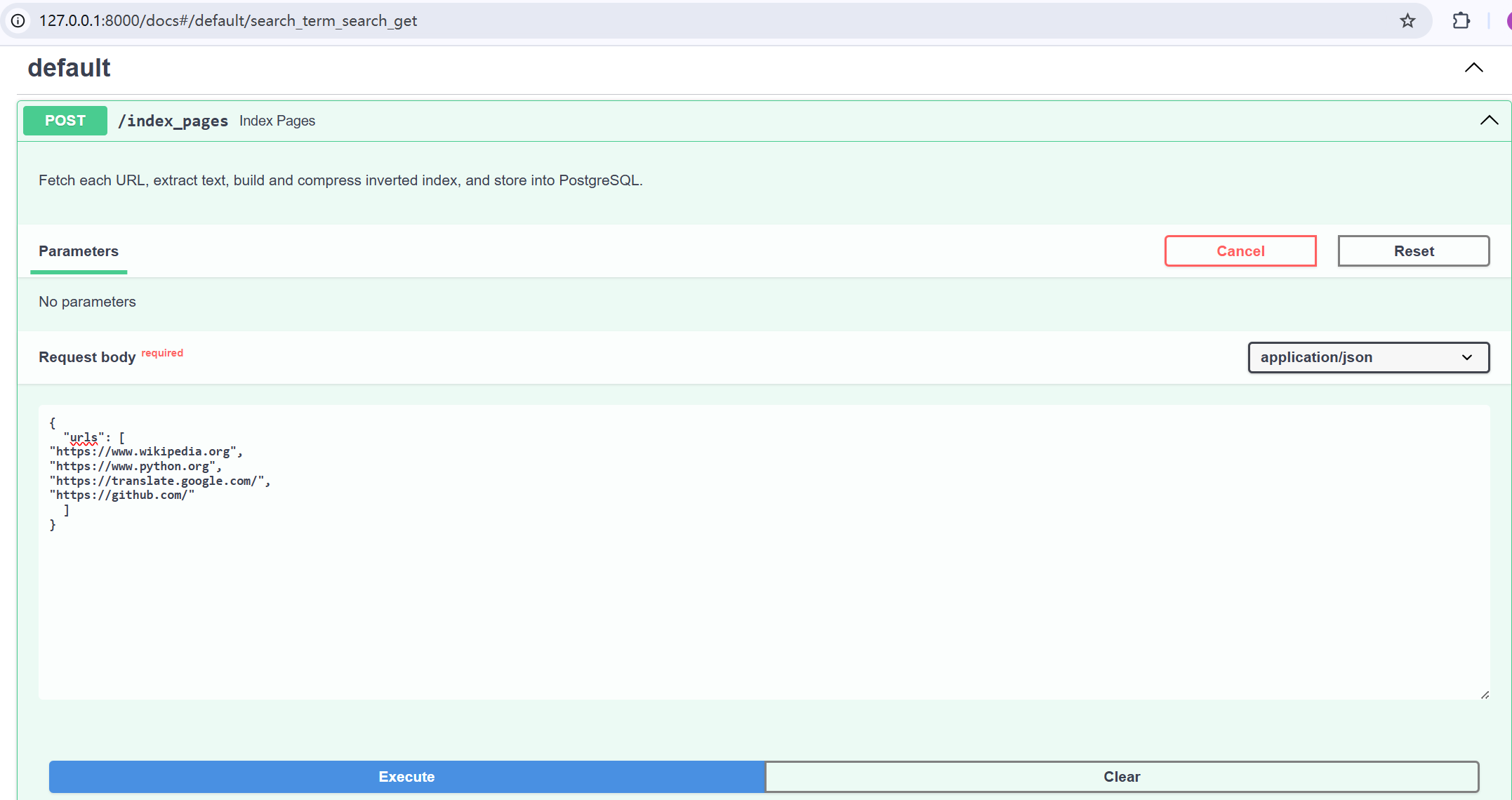
Module 3

**1. Подготовка окружающей среды**

1. **Убедитесь, что база данных PostgreSQL правильно установлена ​​и запущена.**
2. **Создайте базу данных inverted\_index\_db в соответствии с файлом конфигурации (переменные среды или конфигурация в коде), создайте пользователя youruser и предоставьте соответствующие разрешения.**
3. **Установите библиотеки, зависимые от Python (FastAPI, SQLAlchemy, запросы и т. д.).**
4. **Запустите службу FastAPI (с помощью команды uvicorn index\_web\_app:app --reload).**
5. **Посетите страницу документации интерфейса http://127.0.0.1:8000/docs или http://127.0.0.1:8000/redoc, чтобы убедиться, что интерфейс загружается нормально.**
6. **2.测试步骤**

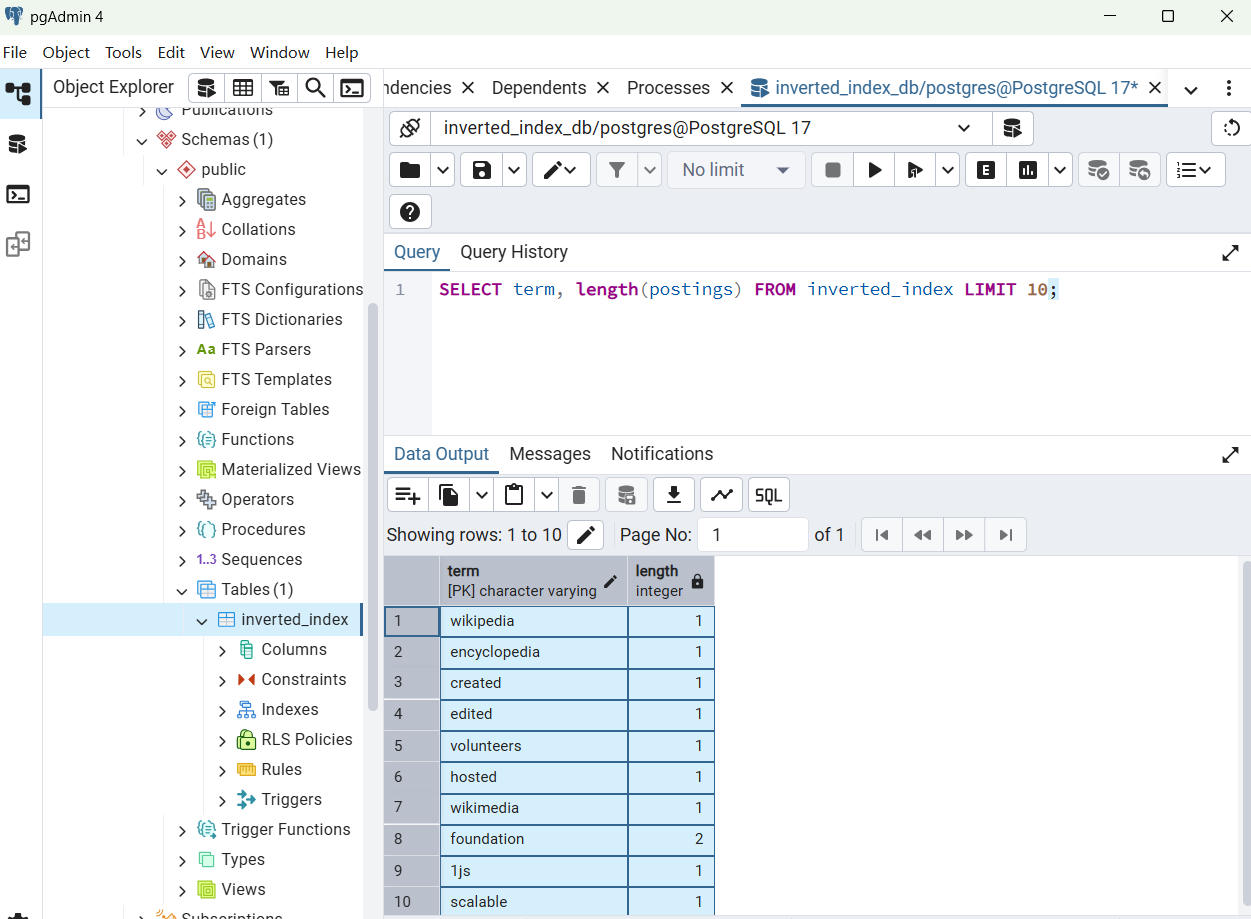
2.1 API-тест index\_pages

Используйте конечную точку POST /index\_pages для отправки запроса, содержащего несколько тестовых URL-адресов.

Убедитесь, что запрос возвращает количество терминов, которые были успешно проиндексированы.

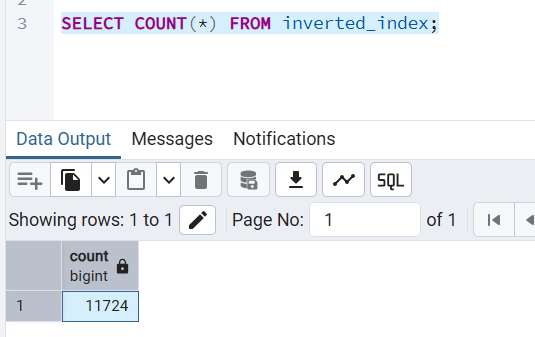
Результат после Execute, код 200, означает, что выполнение прошло успешно:

2.2 Проверка хранилища базы данных

Убедитесь, что сохраненные записи инвертированного индекса существуют и имеют правильный формат (поле записей представляет собой строку байтов), отправив запрос в базу данных.

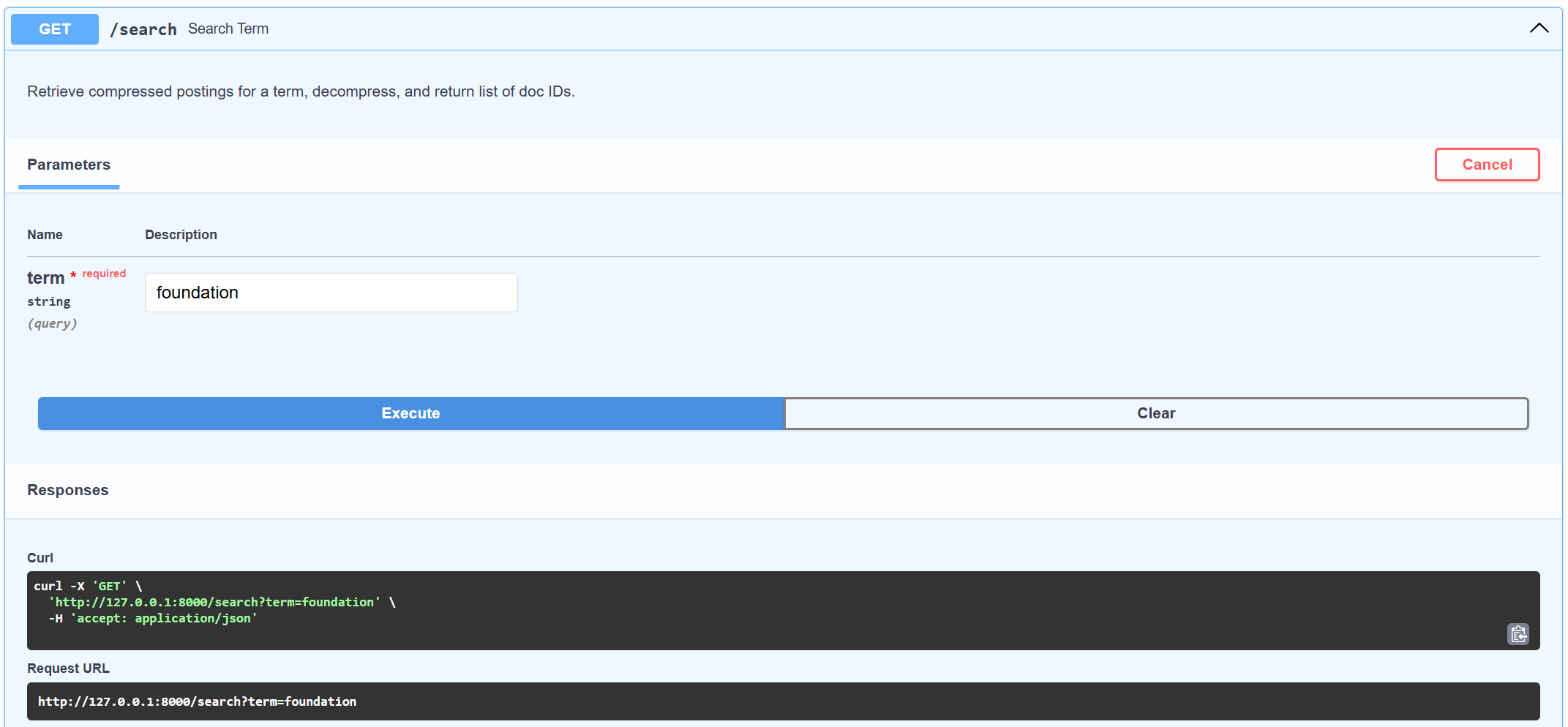
Проверьте, соответствует ли объем данных в базе данных количеству представленных документов и терминов.

В версии 2.1 интерфейс /index\_pages использовался для отправки URL-адресов в общей сложности 4 веб-страниц (таких как Wikipedia, python, github и домашняя страница translate.google).

Выполнить в PostgreSQL: SELECT COUNT(\*) FROM inverted\_index;

**Общее количество терминов составляет 11724, что совпадает с результатом на шаге 2. Это означает, что содержимое веб-страницы было успешно проанализировано и создан инвертированный индекс.**

2.3 /search Тестирование интерфейса

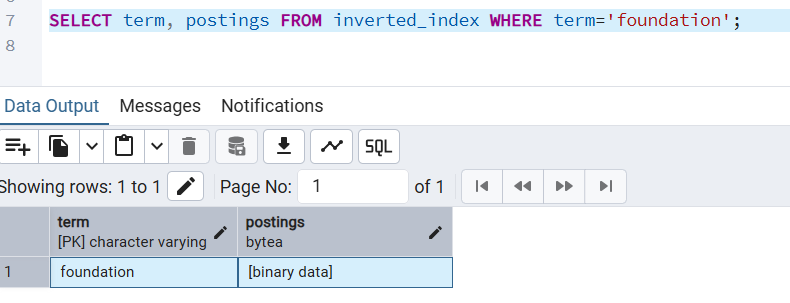
Результаты поиска интерфейса:

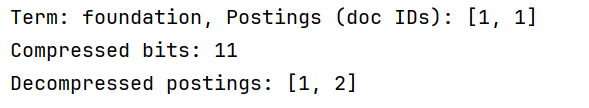


2.4 Проверка содержимого базы данных и сжатого хранилища

Используйте SQL для запроса данных полей записей фонда в таблице inverted\_index базы данных.

ВЫБЕРИТЕ термин, сообщения FROM inverted\_index WHERE term='foundation';

При прямом запросе с помощью SQL вы увидите двоичные данные, а инвертированный список не будет отображаться в виде обычного текста. следующее:

Для декодирования необходимо использовать Python (или другие клиентские программы) совместно с функцией декодирования в проекте. следующее:

Результат декодирования соответствует результату поиска страницы.