Департамент образования и науки города Москвы

Государственное автономное образовательное учреждение

высшего образования города Москвы

«Московский городской педагогический университет»

Институт цифрового образования

Департамент информатики управления и технологий

Кузьмина Дарья Юрьевна БД-241м

Инструменты хранения и анализа больших данных

**Лабораторная работа 2.1. Часть 3.**

**Рекомендательные системы. GraphDB**

**Вариант 11**

Направление подготовки/специальность

38.04.05 - Бизнес-информатика

Бизнес-аналитика и большие данные

(очная форма обучения)

Руководитель дисциплины:

Босенко Т.М., доцент департамента

информатики, управления и технологий,

доктор экономических наук

Москва

2025

Содержание

[Введение 2](#_Toc192845116)

[Основная часть 2](#_Toc192845117)

[Заключение 12](#_Toc192845118)

## Введение

**Цель**

освоение работы с базой данных GraphDB в виртуальной машине с использованием Docker. Студенты научатся загружать RDF-данные в базу, выполнять запросы с использованием SPARQL и анализировать результаты, используя функциональность GraphDB.

**Задачи:**

1. Настроить и запустить контейнер с GraphDB с помощью Docker.
2. Загрузить RDF-данные о фильмах в базу данных.
3. Ознакомиться с основами SPARQL-запросов.
4. Выполнить различные SPARQL-запросы для получения информации из базы данных.
5. Проанализировать и интерпретировать результаты выполнения запросов.

**Необходимое ПО:**

* **Операционная система**: Ubuntu 22.
* **СУБД**: GraphDB.
* **Docker**: для запуска контейнера с GraphDB.
* **Среда разработки**: SPARQL редактор для выполнения запросов.

**Процесс начала работы:**

1. Подключитесь к виртуальной машине и войдите в нее.
2. Перейдите в каталог с проектом:
3. cd ~/Downloads/dba/nonrel/graphdb
4. Остановите контейнер, если он работает:
5. sudo docker compose stop
6. Запустите контейнер с базой данных GraphDB:
7. sudo docker compose start

## Основная часть

Задача 1.

Запускаем GraphDB

Изображение выглядит как текст, Шрифт, снимок экрана

Автоматически созданное описание

Изображение выглядит как текст, программное обеспечение, веб-страница, Значок на компьютере

Автоматически созданное описание

Pаходим в GraphDB

Изображение выглядит как текст, снимок экрана, веб-страница, программное обеспечение

Автоматически созданное описание

Создаем базу данных

Изображение выглядит как текст, программное обеспечение, Шрифт, веб-страница

Автоматически созданное описание

Импортируем данные из источника, тут можно указать свой гит.

После проведения практической части. Я добавила свой датасет на гит:

<https://raw.githubusercontent.com/Iezekiss/BDSAD_MGPU/refs/heads/main/Lab2.1/Lab2.1.3/TMDB_all_movies.csv>

Кстати, там можно посмостреть рейтинг и дату выхода.

Изображение выглядит как текст, снимок экрана, Шрифт

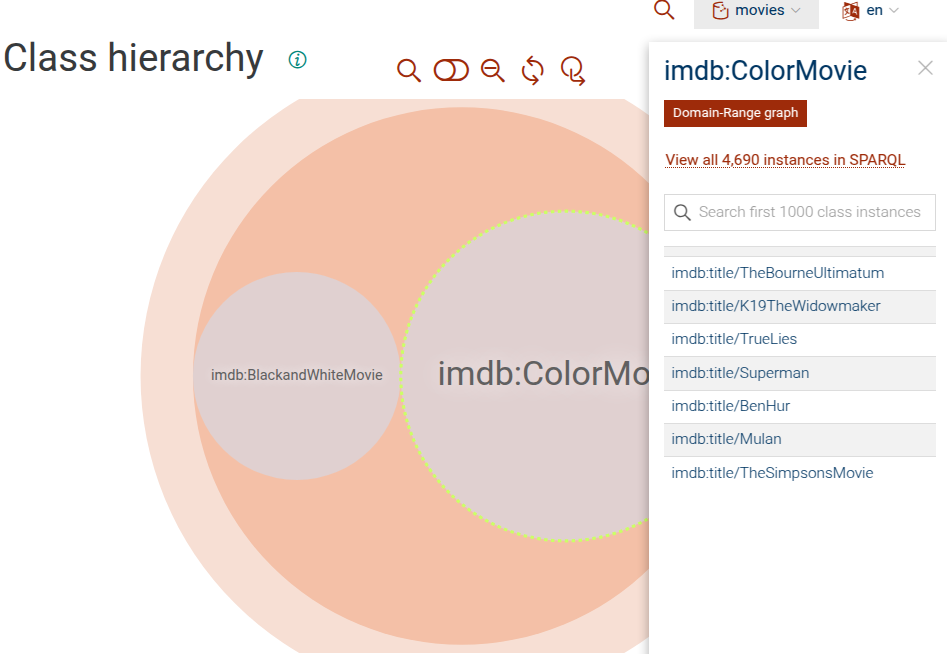
Автоматически созданное описание

Импортировала

Изображение выглядит как текст, снимок экрана, программное обеспечение, Значок на компьютере

Автоматически созданное описание

Смотрим схемы.



Смотрим первичные графы.

Изображение выглядит как текст, снимок экрана, веб-страница, Веб-сайт

Автоматически созданное описание

Посмотрим, какими графами описан объект.

Изображение выглядит как круг, Красочность, диаграмма, снимок экрана

Автоматически созданное описание

Рассмотрим субъект, объект и предикат.

Изображение выглядит как текст, программное обеспечение, веб-страница, Значок на компьютере

Автоматически созданное описание

Переходим в SPARQL.

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Выполняем первый запрос, ограничивающий количество графов до 100

Далее мы нашли набор связанных графов из источников из семантической сети и выполнили дополнительные задания на отработку поисковых запросов.

Индивидуальное задание **Вариант 11**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Напишите запрос для получения всех фильмов с их количеством комментариев. | Создайте запрос для поиска всех фильмов с рейтингом выше 9,0. | Найдите все фильмы, в которых сыграл "Keanu Reeves" и режиссером был "Lana Wachowski". | Напишите запрос для поиска всех фильмов, выпущенных в 1990-е годы. | Найдите все фильмы, у которых актер и режиссер — один и тот же человек. |

Задание 1

**Получение всех фильмов с количеством комментариев**

**Что делает запрос:**

Находит все фильмы и подсчитывает, сколько комментариев есть у каждого. Сортирует по убыванию количества комментариев.

**Пример вывода:**

- Фильм "Матрица" — 150 комментариев

- Фильм "Крестный отец" — 120 комментариев

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Задание 3

**Фильмы с Киану Ривзом в роли актера и Ланой Вачовски в роли режиссера**

**Что делает запрос:**

Ищет фильмы, где одновременно:

- Актер — **"Keanu Reeves"**

- Режиссер — **"Lana Wachowski"**

**Пример вывода:**

- "Матрица"

- "Матрица: Перезагрузка"

Изображение выглядит как текст, снимок экрана, программное обеспечение, веб-страница

Автоматически созданное описание

Задание 5

**Фильмы, где актер и режиссер — один и тот же человек**

**Что делает запрос:**

Ищет фильмы, в которых один человек был и **актером**, и **режиссером**. Показывает название фильма и имя этого человека.

**Пример вывода:**

- "Гражданин Кейн" — Орсон Уэллс

- "Одержимость" — Клинт Иствуд

Изображение выглядит как текст, программное обеспечение, веб-страница, Веб-сайт

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, число, Шрифт

Автоматически созданное описание

Получить данные по заданию 2 и заданию 4 невозможно, поскольку в моем файле нету данных о рейтинге и дате выпуска фильма, однако можно, например, отсортировать по стране выпуска, цветной или не цветной и другие варианты, однако предполагаемые запросы выглядели бы так:

**2. Поиск фильмов с рейтингом выше 9.0**

**Что делает запрос:**

Выбирает фильмы, у которых рейтинг больше 9.0, и показывает их названия и сам рейтинг.

**Пример вывода:**

- "Побег из Шоушенка" — 9.3

- "Крестный отец" — 9.2

Изображение выглядит как текст, программное обеспечение, снимок экрана, веб-страница

Автоматически созданное описание

**4. Фильмы, выпущенные в 1990-е годы**

**Что делает запрос:**

Находит все фильмы, которые вышли с **1990 по 1999 год**, и показывает их названия и год выпуска.

**Пример вывода:**

- "Криминальное чтиво" (1994)

- "Титаник" (1997)

Изображение выглядит как текст, снимок экрана, программное обеспечение, дисплей

Автоматически созданное описание

# Заключение

**Вывод**:

Освоена работа с GraphDB: загрузка данных, настройка репозитория.

Реализованы запросы на SPARQL для анализа связей между сущностями.

Получены навыки фильтрации, агрегации и работы с типами данных.