

# Каталог фильмов «ИнфоМоушен»

Трек: Искусственный интеллект



Команда: Литовская Дарья,  
Токарева Виктория,  
Шейнина Александра

Руководитель: Ворущенко П.В.

# Наша команда



**Литовская  
Дарья**

Разработчик,  
создание  
интерфейса,  
дизайн



**Шейнина  
Александра**

Разработчик,  
создание базы  
данных, создание  
поиска и дизайн  
сайта



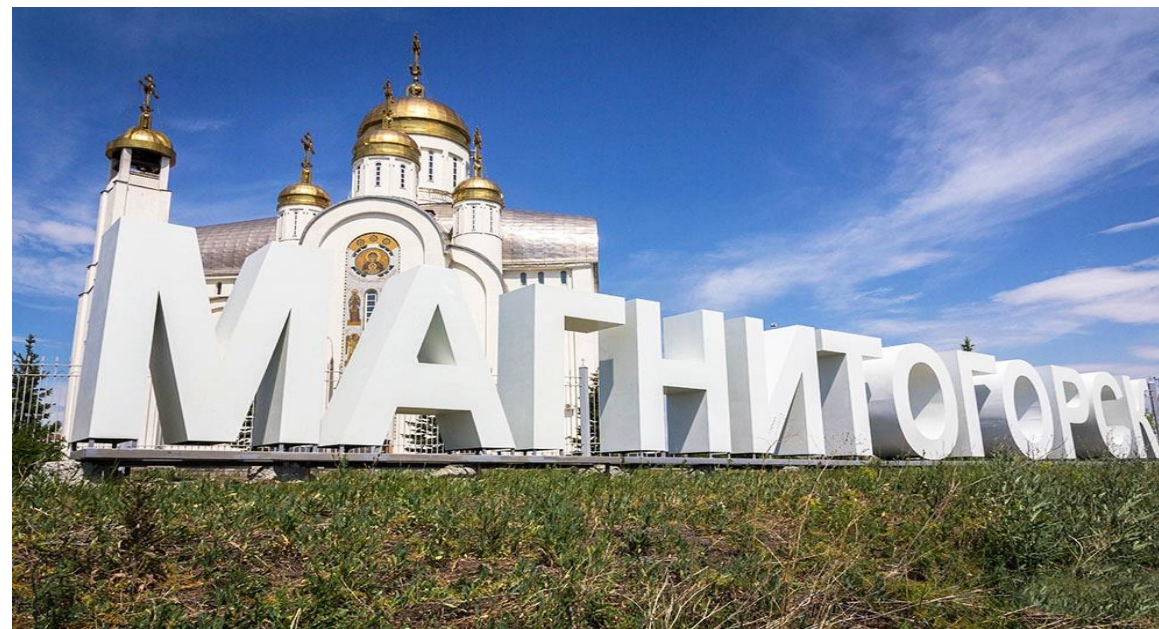
**Токарева  
Виктория**

Разработчик,  
парсинг сайта,  
дизайн



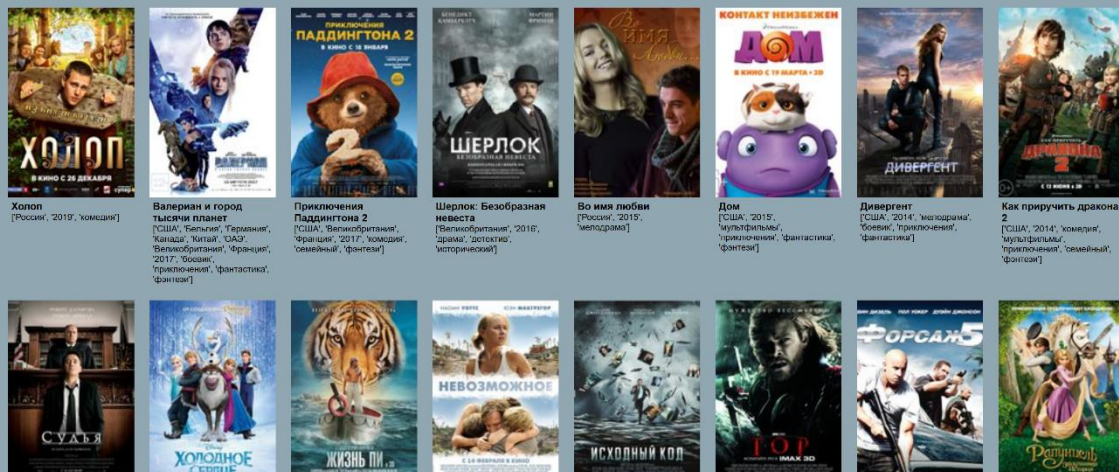
# Наша команда

Наша команда «Инфогении»  
представляет город  
**Магнитогóрск** — город  
в Челябинской области  
являющийся городом трудовой  
доблести, крупным культурным и  
деловым центром Южного Урала,  
одним из крупнейших мировых  
центров чёрной металлургии.





## Подборка фильмов



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

ИНФО  
МОУШЕН

# Цель и задачи

**Цель** - создать веб-сервис с поисковой строкой по собранной базе фильмов.

## Задачи:

- Сделать парсинг сайта с фильмами, собрать необходимую информацию о фильмах;
- Создать базу данных;
- Создать алгоритм поиска;
- Создать сайт с удобным интерфейсом.





#	наим	year	image
1	Человек-паук: Нет пути домой	["США", "2021", "Боевик", ...]	https://cinema.mail.ru/p/7728707/
2	Делители нуля	["США", "Великобритания", "2019", ...]	https://cinema.mail.ru/p/249430a/
3	Давидов	["США", "Триллер", "Боевик", "Действие", ...]	https://cinema.mail.ru/p/60e2b04b/
4	Госа чужеземный	["США", "Великобритания", "2021", ...]	https://cinema.mail.ru/p/...
5	Солдаты	["Россия", "2021", "Драма", ...]	https://cinema.mail.ru/p/...
6	1+1	["Франция", "2021", "Драма", ...]	https://cinema.mail.ru/p/...
7	Вики 2	["США", "2021", "Триллер", "Боевик", ...]	https://cinema.mail.ru/p/...
8	Контрастные твари и как они ...	["США", "Великобритания", "2016", ...]	https://cinema.mail.ru/p/...
9	Неролика	["США", "Великобритания", "2019", ...]	https://cinema.mail.ru/p/...
10	Лига справедливости Зима Снайпера	["США", "2021", "Боевик", ...]	https://cinema.mail.ru/p/7767b04b/
11	Небо	["Россия", "2021", "Драма", ...]	https://cinema.mail.ru/p/6af3b07b/
12	Мстители: Финал	["США", "2019", "Боевик", ...]	https://cinema.mail.ru/p/69b450c/
13	Долгая ночь	["США", "2019", "Драма", "Боевик", ...]	https://cinema.mail.ru/p/28537e4b/
14	Скиталец	["Россия", "2018", "Драма", ...]	https://cinema.mail.ru/p/...
15	Собачья жизнь 2	["США", "Россия", "Драма", "Китай", ...]	https://cinema.mail.ru/p/...
16	Зеленая книга	["США", "2018", "Драма", ...]	https://cinema.mail.ru/p/...
17	Мстители: Война бесконечности	["США", "2018", "Боевик", ...]	https://cinema.mail.ru/p/...
18	Давидов: Конец	["Россия", "2017", "Драма", "Сюжет", ...]	https://cinema.mail.ru/p/...
19	Три: Ретардер	["США", "2017", "Боевик", ...]	https://cinema.mail.ru/p/...
20	До встречи с тобой	["США", "Великобритания", "2016", ...]	https://cinema.mail.ru/p/20140a0b/
21	Мстители: Ура Льюиса	["США", "2019", "Боевик", ...]	https://cinema.mail.ru/p/...
22	Ворон 7	["США", "Бразилия", "2015", "Триллер", ...]	https://cinema.mail.ru/p/5e7488f/
23	Зеленая книга	["США", "1990", "Драма", "Детектив", ...]	https://cinema.mail.ru/p/...
24	Бойер из Юта	["США", "1994", "Драма", "Криминал", ...]	https://cinema.mail.ru/p/...
25	Форд против Феттера	["США", "2019", "Драма", "Биография", ...]	https://cinema.mail.ru/p/...
26	Доктор	["США", "2018", "Драма", "Триллер", ...]	https://cinema.mail.ru/p/...
27	Богемская рапсодия	["США", "Великобритания", "2018", ...]	https://cinema.mail.ru/p/...
28	Время первого	["Россия", "2017", "Триллер", ...]	https://cinema.mail.ru/p/...
29	Дело Уин 2	["США", "2017", "Триллер", "Боевик", ...]	https://cinema.mail.ru/p/...
30	Логан	["США", "2017", "Драма", "Триллер", ...]	https://cinema.mail.ru/p/...
31	Половина богатыря	["Россия", "2017", "Боевик", ...]	https://cinema.mail.ru/p/...

```

import requests
from bs4 import BeautifulSoup
url = 'https://kino.mail.ru/c'
baseurl='https://kino.mail.ru'
responce = requests.get(url)
html = requests.get(url)

```

```

c_light_large
<a class="link-holder link-holder_itemevent link-holder_itemevent_small" href="/cinema/movie/925247_spider_man_no_way_home/">
::before == $0
<span class="link_text">Человек-паук: Нет пути домой</span>
</a>
</span>
<span class="text text_block text_fixed t_light_small color_gray margin_bottom_5">Spider-Man: No Way Home</span>
</div>

```

```

link_holder.link-holder.link-holder_itemevent.link-holder_itemevent_small
::before
и Вычисляемые Макет Прослушиватели событий Точки останова DOM
Фильтр :hov .cls +
.link-holder_itemevent_small:before {
common.cs:
icht 165px

```

# Сбор базы данных

Наша команда выполнила Парсинг сайта Кино Mail и собрала базу данных о 1000 фильмах с сайта [«Кино Mail»](#) с помощью библиотек requests и BeautifulSoup. В итоге мы создали базу фильмов с помощью класса models из модуля Django.db

# Способ сбора базы данных

Для сбора базы фильмов мы воспользовались **Парсингом** сайта [Kino.mail.ru](http://Kino.mail.ru) (парсинг сайтов - это процесс автоматического извлечения данных с веб-страниц). Для этого мы создали функцию в файле [Parser.py](#), которая сканирует исходный HTML-код и вычленяет семантически значимые единицы по заданным параметрам, после чего добавляет нужные значения в базу данных.

DB Browser for SQLite - C:\Users\sheynina135599\Desktop\kinopoisk\db.sqlite3

Файл Редактирование Вид Инструменты Справка

New Database Open Database Записать изменения Отменить изменения Undo Open Project Save Project Прикрепить БД Закрывать

Database Structure Browse Data Edit Pragma Execute SQL

Таблица: main\_films

	id	name	teges	image
Фил...	Фильтр	Фильтр	Фильтр	Фильтр
1	1	Человек-паук: Нет пути домой	['США', '2021', 'боевик', ...]	/films/code_geass1.jpg
2	2	Джентльмены	['США', 'Великобритания', '2019', ...]	/films/code_geass2.jpg
3	3	Выживший	['США', 'Гонконг', 'Канада', 'Тайван...]	/films/code_geass3.jpg
4	4	Гнев человеческий	['США', 'Великобритания', '2021', ...]	/films/code_geass4.jpg
5	5	Солнцепек	['Россия', '2021', 'драма']	/films/code_geass5.jpg
6	6	1+1	['Франция', '2011', 'драма', ...]	/films/code_geass6.jpg
7	7	Веном 2	['США', '2021', 'триллер', 'боевик', ...]	/films/code_geass7.jpg
8	8	Фантастические твари и где они ...	['США', 'Великобритания', '2016', ...]	/films/code_geass8.jpg
9	9	Марсианин	['США', 'Великобритания', '2015', ...]	/films/code_geass9.jpg
10	10	Лига справедливости Зака Снайдера	['США', '2021', 'боевик', ...]	/films/code_geass10.jpg
11	11	Небо	['Россия', '2021', 'драма', ...]	/films/code_geass11.jpg
12	12	Мстители: Финал	['США', '2019', 'боевик', ...]	/films/code_geass12.jpg
13	13	Зеленая книга	['США', '2018', 'драма']	/films/code_geass13.jpg
14	14	Достать ножи	['США', '2019', 'драма', 'комедия', ...]	/films/code_geass14.jpg
15	15	Сестренка	['Россия', '2019', 'драма', ...]	/films/code_geass15.jpg
16	16	Собачья жизнь 2	['США', 'Гонконг', 'Индия', 'Китай', ...]	/films/code_geass16.jpg
17	17	Мстители: Война бесконечности	['США', '2018', 'боевик', ...]	/films/code_geass17.jpg
18	18	Движение вверх	['Россия', '2017', 'драма', 'спорт']	/films/code_geass18.jpg
19	19	Тор: Рагнарек	['США', '2017', 'боевик', ...]	/films/code_geass19.jpg
20	20	До встречи с тобой	['США', 'Великобритания', '2016', ...]	/films/code_geass20.jpg
21	21	Мстители: Эра Альтрона	['США', '2015', 'боевик', ...]	/films/code_geass21.jpg
22	22	Форсаж 7	['США', 'Япония', '2015', 'триллер', ...]	/films/code_geass22.jpg
23	23	Зеленая миля	['США', '1999', 'драма', 'детектив', ...]	/films/code_geass23.jpg
24	24	Побег из Шоушенка	['США', '1994', 'драма', 'криминал']	/films/code_geass24.jpg
25	25	Ford против Ferrari	['США', '2019', 'драма', 'биография', ...]	/films/code_geass25.jpg
26	26	Джокер	['США', '2019', 'драма', 'триллер', ...]	/films/code_geass26.jpg
27	27	Богемская рапсодия	['США', 'Великобритания', '2018', ...]	/films/code_geass27.jpg
28	28	Время первых	['Россия', '2017', 'триллер', ...]	/films/code_geass28.jpg
29	29	Джон Уик 2	['США', '2017', 'триллер', 'боевик', ...]	/films/code_geass29.jpg
30	30	Логан	['США', '2017', 'драма', 'триллер', ...]	/films/code_geass30.jpg
31	31	Последний богатырь	['Россия', '2017', 'комедия', ...]	/films/code_geass31.jpg
32	32	Стражи Галактики. Часть 2	['США', '2017', 'комедия', 'боевик', ...]	/films/code_geass32.jpg

1 - 32 of 1 000

Перейти к: 1



# Способ сбора базы данных

Мы воспользовались библиотеками requests и BeautifulSoup(bs4). Библиотека requests нужна для упрощения работы с HTTP-запросами. Она позволяет отправлять запросы веб-сервисам и получать соответствующие ответы. BeautifulSoup создает дерево из HTML-кода, что позволяет легко получить доступ к различным элементам веб-страницы. Парсинг осуществлялся по ссылке <https://kino.mail.ru/cinema/all/>



DB Browser for SQLite - C:\Users\sheynina135599\Desktop\kinopoisk\db.sqlite3

Файл Редактирование Вид Инструменты Справка

New Database Open Database Записать изменения Отменить изменения Undo Open Project Save Project Прикрепить БД Закр...

Database Structure Browse Data Edit Pragma Execute SQL

Таблица: main\_films

	id	name	teges	image
Фил...	Фильтр		Фильтр	Фильтр
1	1	Человек-паук: Нет пути домой	['США', '2021', 'боевик', ...]	/films/code_geass1.jpg
2	2	Джентльмены	['США', 'Великобритания', '2019', ...]	/films/code_geass2.jpg
3	3	Выживший	['США', 'Гонконг', 'Канада', 'Тайван...]	/films/code_geass3.jpg
4	4	Гнев человеческий	['США', 'Великобритания', '2021', ...]	/films/code_geass4.jpg
5	5	Солнцепек	['Россия', '2021', 'драма']	/films/code_geass5.jpg
6	6	1+1	['Франция', '2011', 'драма', ...]	/films/code_geass6.jpg
7	7	Веном 2	['США', '2021', 'триллер', 'боевик', ...]	/films/code_geass7.jpg
8	8	Фантастические твари и где они ...	['США', 'Великобритания', '2016', ...]	/films/code_geass8.jpg
9	9	Марсианин	['США', 'Великобритания', '2015', ...]	/films/code_geass9.jpg
10	10	Лига справедливости Зака Снайдера	['США', '2021', 'боевик', ...]	/films/code_geass10.jpg
11	11	Небо	['Россия', '2021', 'драма', ...]	/films/code_geass11.jpg
12	12	Мстители: Финал	['США', '2019', 'боевик', ...]	/films/code_geass12.jpg
13	13	Зеленая книга	['США', '2018', 'драма']	/films/code_geass13.jpg
14	14	Достать ножи	['США', '2019', 'драма', 'комедия', ...]	/films/code_geass14.jpg
15	15	Сестренка	['Россия', '2019', 'драма', ...]	/films/code_geass15.jpg
16	16	Собачья жизнь 2	['США', 'Гонконг', 'Индия', 'Китай', ...]	/films/code_geass16.jpg
17	17	Мстители: Война бесконечности	['США', '2018', 'боевик', ...]	/films/code_geass17.jpg
18	18	Движение вверх	['Россия', '2017', 'драма', 'спорт']	/films/code_geass18.jpg
19	19	Тор: Рагнарек	['США', '2017', 'боевик', ...]	/films/code_geass19.jpg
20	20	До встречи с тобой	['США', 'Великобритания', '2016', ...]	/films/code_geass20.jpg
21	21	Мстители: Эра Альтрона	['США', '2015', 'боевик', ...]	/films/code_geass21.jpg
22	22	Форсаж 7	['США', 'Япония', '2015', 'триллер', ...]	/films/code_geass22.jpg
23	23	Зеленая миля	['США', '1999', 'драма', 'детектив', ...]	/films/code_geass23.jpg
24	24	Побег из Шоушенка	['США', '1994', 'драма', 'криминал']	/films/code_geass24.jpg
25	25	Ford против Ferrari	['США', '2019', 'драма', 'биография', ...]	/films/code_geass25.jpg
26	26	Джокер	['США', '2019', 'драма', 'триллер', ...]	/films/code_geass26.jpg
27	27	Богемская рапсодия	['США', 'Великобритания', '2018', ...]	/films/code_geass27.jpg
28	28	Время первых	['Россия', '2017', 'триллер', ...]	/films/code_geass28.jpg
29	29	Джон Уик 2	['США', '2017', 'триллер', 'боевик', ...]	/films/code_geass29.jpg
30	30	Логан	['США', '2017', 'драма', 'триллер', ...]	/films/code_geass30.jpg
31	31	Последний богатырь	['Россия', '2017', 'комедия', ...]	/films/code_geass31.jpg
32	32	Стражи Галактики. Часть 2	['США', '2017', 'комедия', 'боевик', ...]	/films/code_geass32.jpg

1 - 32 of 1 000

Перейти к: 1

# Способ сбора базы данных

DB Browser for SQLite - C:\Users\sheynina135599\Desktop\kinopoisk\db.sqlite3

Файл Редактирование Вид Инструменты Справка

New Database Open Database Записать изменения Отменить изменения Undo Open Project Save Project Прикрепить БД Закрывать базу данных

Database Structure Browse Data Edit Pragmas Execute SQL

Таблица: main\_films

	id	name	teges	image
Фил...	Фильтр	Фильтр	Фильтр	Фильтр
1	1	Человек-паук: Нет пути домой	['США', '2021', 'боевик', ...]	/films/code_geass1.jpg
2	2	Джентльмены	['США', 'Великобритания', '2019', ...]	/films/code_geass2.jpg
3	3	Выживший	['США', 'Гонконг', 'Канада', 'Тайван...	/films/code_geass3.jpg
4	4	Гнев человеческий	['США', 'Великобритания', '2021', ...]	/films/code_geass4.jpg
5	5	Солнцекек	['Россия', '2021', 'драма']	/films/code_geass5.jpg
6	6	1+1	['Франция', '2011', 'драма', ...]	/films/code_geass6.jpg
7	7	Веном 2	['США', '2021', 'триллер', 'боевик', ...]	/films/code_geass7.jpg
8	8	Фантастические твари и где они ...	['США', 'Великобритания', '2016', ...]	/films/code_geass8.jpg
9	9	Марсианин	['США', 'Великобритания', '2015', ...]	/films/code_geass9.jpg
10	10	Лига справедливости Зака Снайдера	['США', '2021', 'боевик', ...]	/films/code_geass10.jpg
11	11	Небо	['Россия', '2021', 'драма', ...]	/films/code_geass11.jpg
12	12	Мстители: Финал	['США', '2019', 'боевик', ...]	/films/code_geass12.jpg
13	13	Зеленая книга	['США', '2018', 'драма']	/films/code_geass13.jpg
14	14	Достать ножи	['США', '2019', 'драма', 'комедия', ...]	/films/code_geass14.jpg
15	15	Сестренка	['Россия', '2019', 'драма', ...]	/films/code_geass15.jpg
16	16	Собачья жизнь 2	['США', 'Гонконг', 'Индия', 'Китай', ...]	/films/code_geass16.jpg
17	17	Мстители: Война бесконечности	['США', '2018', 'боевик', ...]	/films/code_geass17.jpg
18	18	Движение вверх	['Россия', '2017', 'драма', 'спорт']	/films/code_geass18.jpg
19	19	Тор: Рагнарек	['США', '2017', 'боевик', ...]	/films/code_geass19.jpg
20	20	До встречи с тобой	['США', 'Великобритания', '2016', ...]	/films/code_geass20.jpg
21	21	Мстители: Эра Альтрона	['США', '2015', 'боевик', ...]	/films/code_geass21.jpg
22	22	Форсаж 7	['США', 'Япония', '2015', 'триллер', ...]	/films/code_geass22.jpg
23	23	Зеленая миля	['США', '1999', 'драма', 'детектив', ...]	/films/code_geass23.jpg
24	24	Побег из Шоушенка	['США', '1994', 'драма', 'криминал']	/films/code_geass24.jpg
25	25	Ford против Ferrari	['США', '2019', 'драма', 'биография', ...]	/films/code_geass25.jpg
26	26	Джокер	['США', '2019', 'драма', 'триллер', ...]	/films/code_geass26.jpg
27	27	Вогемская рапсодия	['США', 'Великобритания', '2018', ...]	/films/code_geass27.jpg
28	28	Время первых	['Россия', '2017', 'триллер', ...]	/films/code_geass28.jpg
29	29	Джон Уик 2	['США', '2017', 'триллер', 'боевик', ...]	/films/code_geass29.jpg
30	30	Логан	['США', '2017', 'драма', 'триллер', ...]	/films/code_geass30.jpg
31	31	Последний богатырь	['Россия', '2017', 'комедия', ...]	/films/code_geass31.jpg
32	32	Стражи Галактики. Часть 2	['США', '2017', 'комедия', 'боевик', ...]	/films/code_geass32.jpg

1 - 32 of 1 000

Перейти к: 1

Редактирование ячейки БД

Режим: Текст

1

Editing row=1, column=0  
Type: Text / Numeric; Size: 1 character(s)

Удаленный сервер

ID Select an identity to connect

DBHub.io Local Current Database

Имя	Изменен	Размер	Коммит
-----	---------	--------	--------

Журнал SQL График Схема БД Удаленный сервер

UTF-8



# Алгоритм поиска

Мы в файле base.html прописали **<form action="{% url 'search' %}"** для определения адреса, куда отправляется запрос. В [urls.py](#) страница search указывает на функцию search, находящуюся в файле view. Эта функция сравнивает запрос пользователя с данными в базе (не считая при этом служебные слова(и, или))

```
6 urlpatterns = [  
7     path('', views.index, name='home'),  
8     path('about', views.about, name='about'),  
9     path('search', views.search, name='search'),  
10 ]
```

```
<form action="{% url 'search' %}" method="POST" class="d-flex editContent">  
    {% csrf_token %}  
    <input type="text" placeholder="Поиск" name="name" class="form-control" required="">  
    <button type="submit" class="search-box">Найти</button>  
</form>
```



# Алгоритм поиска

В итоге функция возвращает значение, которое соответствует запросу пользователя, если оно имеется в базе данных (это может быть название фильма или тег). Значение передаются обратно в base.html, который отвечает за вывод нужного результата

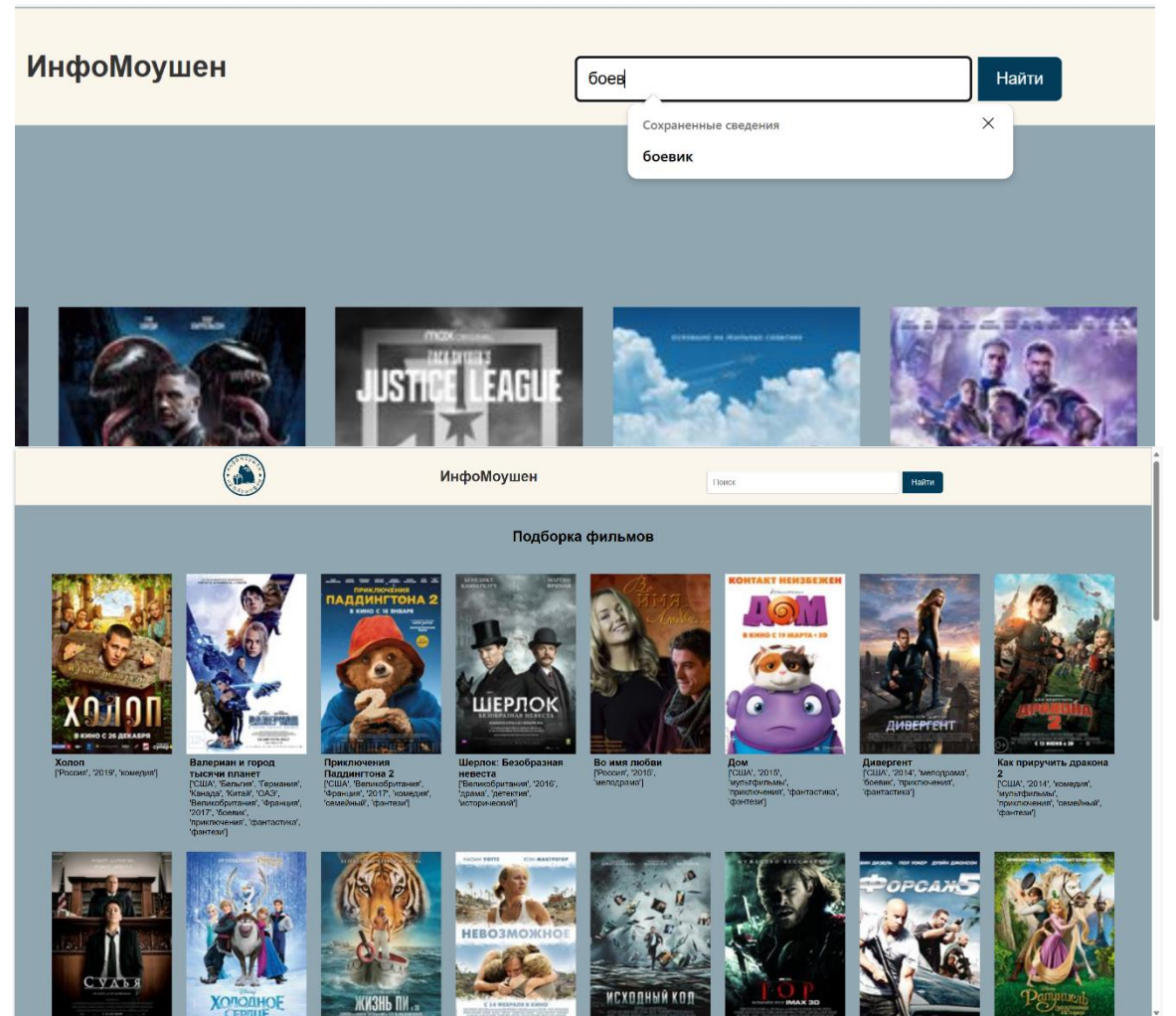
```
31  def search(request): 1 usage
32  if request.method == 'POST':
33      query_name = request.POST.get('name', None)
34  if query_name:
35      stop_words = ['и', 'или', 'в', 'над', 'с', 'без']
36      search_terms = []
37      for term in query_name.split():
38          if term.lower() not in stop_words:
39              if term[-1] != ',':
40                  search_terms.append(term)
41              else:
42                  search_terms.append(term[:len(term)-1])
43      q_objects = Q()
44      for term in search_terms:
45          q_objects &= Q(teges__icontains=term) | Q(name__icontains=term)
46      results = Films.objects.filter(q_objects)
47      return render(request, template_name='main/base.html', context={'results': results})
48  return render(request, template_name='main/base.html')
```





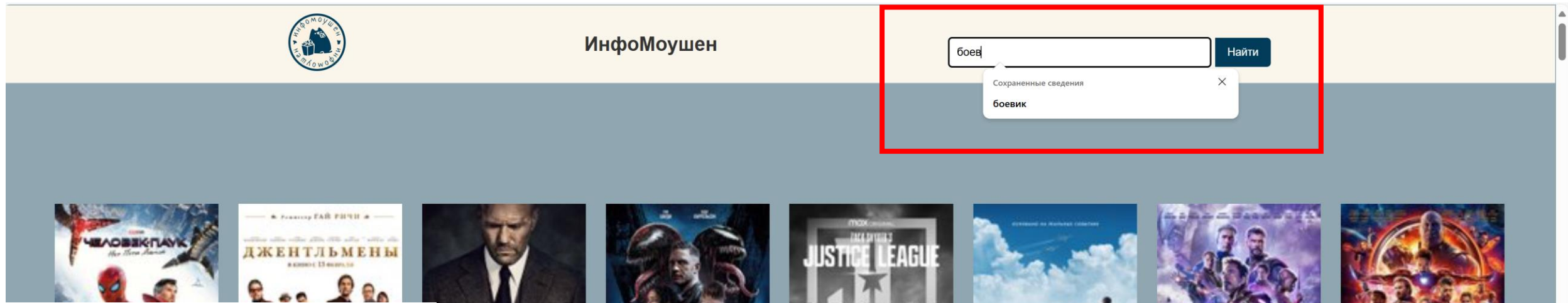
# Веб сервис

Наш веб - сервис «ИнфоМоушен» создан с помощью HTML, CSS, Python с подключением Django. На веб-сервисе представлен список фильмов из базы данных, которую мы создали. На каждой странице представлено 48 фильмов, имеется возможность перелистывания страниц от первой до последней страницы. Веб - сервис осуществляет поиск фильма по названию и тегам



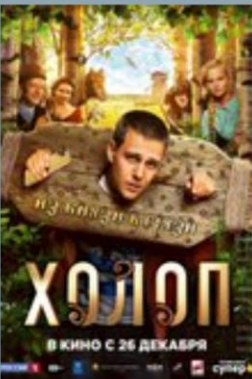


# Веб - сервис





## Подборка фильмов



**Холоп**  
[Россия, '2019', 'комедия']



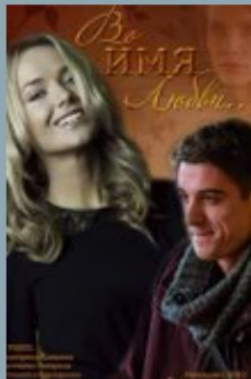
**Валериан и город тысячи планет**  
[США, 'Бельгия', 'Германия', 'Канада', 'Китай', 'ОАЭ', 'Великобритания', 'Франция', '2017', 'боевик', 'приключения', 'фантастика', 'фэнтези']



**Приключения Паддингтона 2**  
[США, 'Великобритания', 'Франция', '2017', 'комедия', 'семейный', 'фэнтези']



**Шерлок: Безобразная невеста**  
[Великобритания, '2016', 'драма', 'детектив', 'исторический']



**Во имя любви**  
[Россия, '2015', 'мелодрама']



**Дом**  
[США, '2015', 'мультфильмы', 'приключения', 'фантастика', 'фэнтези']



**Дивергент**  
[США, '2014', 'мелодрама', 'боевик', 'приключения', 'фантастика']



**Как приручить дракона 2**  
[США, '2014', 'комедия', 'мультфильмы', 'приключения', 'семейный', 'фэнтези']





# Веб - сервис

'приключения', 'фантастика']



**Одна**  
[Россия, '2022', 'драма',  
'триллер', 'боевик',  
'приключения']

['США', '1981', 'боевик',  
'приключения']



**Лед 2**  
[Россия, '2020', 'драма',  
'мелодрама', 'мюзикл',  
'спорт']



**Оно 2**  
[США, 'Канада', '2019',  
'ужасы']



**Дуэлянт**  
[Россия, '2016', 'драма',  
'триллер']



**Завтрак у папы**  
[Россия, '2016', 'комедия']



**Книга джунглей**  
[США, 'Великобритания',  
'2016', 'драма',  
'приключения', 'семейный',  
'фэнтези']

'фэнтези']



**Коробка**  
[Россия, '2016', 'драма',  
'комедия', 'спорт']



**007: СПЕКТР**  
[США, 'Великобритания',  
'2015', 'триллер', 'боевик',  
'приключения']

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

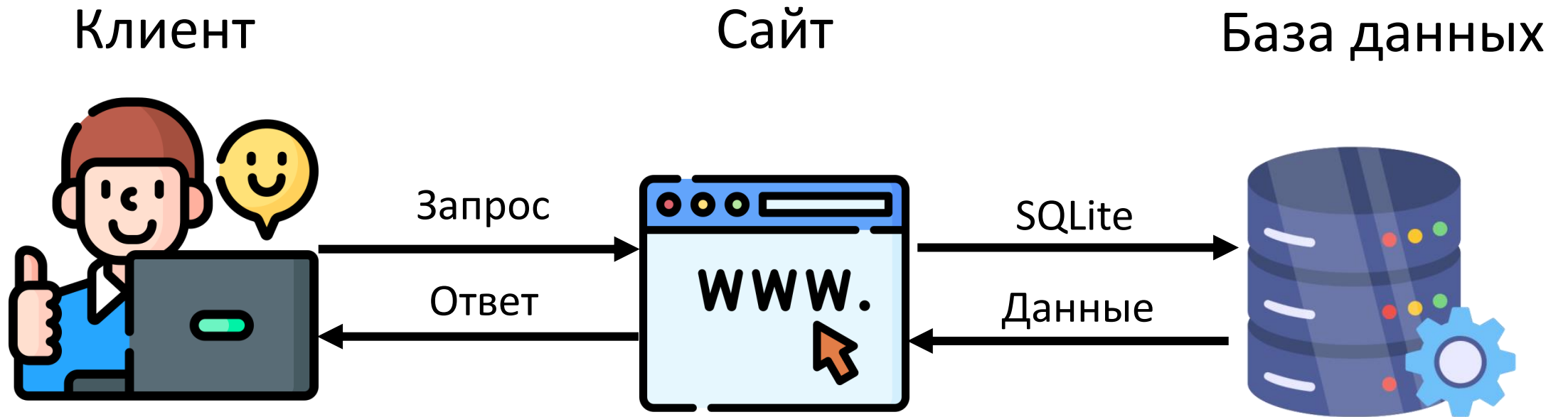


# ИНФО МОУШЕН





# Архитектура реализуемого проекта



**Спасибо за  
внимание!**

