

Разработать приложение для парсинга страниц.

Требования:

Серверная часть приложения должна быть написана на языке Python;

В процессе парсинга страница не должна перезагружаться;

Сценарий работы:

Пользователь открывает страницу.

Вводит список URL до 5 штук за раз. Далее вводит информацию о дате времени, иначе используется текущая дата. Нажимает на кнопку Ок.

Затем в режиме реального времени (без перезагрузки страницы), в специальных блоках наблюдает за прогрессом, статусом парсинга, который начинается в указанное время. Процесс может быть остановлен спец кнопкой.

Это должен быть правдивый прогресс бар или меняющийся список статусов.

Парсинг всех введенных URL должен происходить параллельно. Если имеются 5 URL которые еще не подверглись парсингу, то создавать новые задачи нельзя.

Результатом парсинга одного URL является HTML блок который ранее содержал статус. После завершения процесса он должен содержать:

1. URL
2. Содержимое тега title
3. Содержимое первого тега H1, если он есть
4. Первое изображение из тега img, если он есть. Оно должно быть на фоне этого блока (ссылка на копию изображения на сервисе)

Изображение должно быть закачено на сервис и процесс загрузки этого файла нужно отобразить в прогресс баре.

Обновление страницы не должно приводить к потере каких либо данных или прогресса парсинга.

Блоки с результатами парсинга должны быть разбиты на станицы по 3 блока на одной

Желательно оформить сервис в систему виртуализации docker-compose, Vagrant и пр...