

Тестовое задание на позицию Junior Backend Developer (Java)

Требуется:

Создать приложение для скачивания новостных статей

Источник:

В качестве источника новостей использовать REST API: https://www.spaceflightnewsapi.net/documentation#/Article/get_api_v2_articles

Описание:

Данное описание является исчерпывающим и не требует каких-либо уточнений. Если в процессе выполнения задания у соискателя появляются какие-либо вопросы, просьба решить их по своему собственному усмотрению.

Для выполнения задания потребуется использование REST API:

Метод	URL	Описание
GET	https://test.spaceflightnewsapi.net/api/v2/articles?_limit={articles_limit}&_start={skipped_articles}	<p>Загружает информацию о новостных статьях.</p> <p>Используемые параметры:</p> <ul style="list-style-type: none">• <code>_limit</code> - лимит на количество одновременно загружаемых записей• <code>_start</code> - сколько записей будет пропущено начиная с первой <p>Пример ответа:</p> <pre>[{ "id": "6007d5cc858d39001c60635e", "title": "Video: SLS core stage test-firing ends early", "url": "https://spaceflightnow.com/2021/01/16/video-sls-core-stage-test-firing-ends-early/", "imageUrl": "https://mk0spaceflightnoa02a.kinstacdn.com/wp-content/uploads/test.jpg", "newsSite": "Spaceflight Now", "summary": "The four RS-25 engines at the bottom of NASA's first....", "publishedAt": "2021-01-16T07:03:39.000Z", "updatedAt": "2021-01-20T07:03:40.102Z" }]</pre>

Требуется создать приложение, выполняющее следующие функции:

Реализовать Скачивание новостей:

- Для новостных статей должна быть подготовлена таблица **ARTICLES** в базе данных:

id	title	news_site	published_date	article
Идентификатор статьи	Название статьи	Новостной сайт	Дата публикации	Содержимое статьи

- При старте приложения должен создаваться пул потоков (количество потоков должно быть настраиваемым). Потоки из данного пула должны параллельно скачивать информацию о новостных статьях (использовать API, описанное выше). Общее количество скачиваемых записей, а также количество записей, скачиваемых одним потоком за один цикл работы, должно быть настраиваемым.
- Создать настраиваемый "Черный список новостей". Он должен представлять из себя список слов. Если какое-либо из слов входит в название статьи (поле **title**), то эта статья должна быть исключена из дальнейшей обработки.
- Полученные записи должны быть отсортированы по дате публикации (поле **publishedAt**).
- После сортировки записи должны быть сгруппированы по названию новостного сайта (поле **newsSite**).
- Создать буфер для накопления сгруппированных статей. Буфер должен быть общим для всего пула потоков.

- Каждый раз, когда поток помещает записи в буфер, должна происходить проверка количества записей в буфере для каждого новостного сайта. Если количество записей в буфере больше или равно заданного (настраиваемого) лимита, то необходимо:
 - взять все записи для данного новостного сайта
 - скачать содержимое статей, используя поле **url**
 - поместить всю необходимую информацию для скачанных статей в таблицу **ARTICLES**
 - записи о скачанных статьях должны быть удалены из буфера

Данный механизм должен запускаться в случае, если скачаны все необходимые записи, независимо от достижения лимита.

Реализовать API по выдаче новостных статей:

Реализовать HTTP API по выдаче новостных статей, который :

- Выдает все сохраненные в БД статьи
- Даёт возможность запросить конкретную статью по идентификатору
- Даёт возможность запросить список статей по новостному сайту

Используемые технологии и прочие требования:

- Java 15
- Любые инструменты из экосистемы Spring (Spring boot, Spring MVC и т.д.)
- Хостинг исходного кода Github. Приложение должно быть доступно по прямой ссылке
- Легковесная SQL БД, рекомендуется использовать H2
- Использование прочих библиотек и фреймворков не возбраняется
- Приложение должно компилироваться и исполняться без ошибок, выполняя свои функции в соответствии с заданием