

Индивидуальное задание на курсовое проектирование

Тема: «Разработка многопоточного REST API Scraper»

Постановка задачи:

Необходимо выбрать не менее трёх открытых, периодически обновляемых REST API, например, открытые источники новостей, статей, погодных наблюдений или API различных сервисов, таких как соц. сети, онлайн кинотеатры и т. д. (списки открытых API - [раз](#), [два](#), [три](#))

Следующим шагом нужно разработать приложение, которое в нескольких потоках опрашивает API источников данных и сохраняет новые записи в файл формата CSV или JSON (по выбору пользователя)

Требования:

- 1) В качестве аргумента при запуске приложения передаётся:
 - Количество допустимых одновременно активных потоков n ;
 - Временной таймаут t в секундах, задающий перерыв между итерациями опроса сервиса;
 - Список названий сервисов, которые будут опрашиваться;
 - Формат файла для сохранения результатов
- 2) На каждый источник данных создаётся отдельный поток, в котором производится опрос API, а именно выполнение HTTP запроса к выбранным конечным точкам, возвращающим данные. Одновременно могут работать не более n таких потоков. Если число опрашиваемых сервисов больше n , то потоки для оставшихся сервисов ждут очереди, пока один из других потоков не завершит обновление данных. При этом поток опроса запускается по новой через заданное время t после завершения и также ожидает очереди, если число активных потоков опроса равно n .
- 3) Все потоки пишут данные в один и тот же файл
- 4) Для создания потоков использовать библиотеку `util.concurrent`
- 5) Код должен быть покрыт unit-тестам не менее чем на 70%.
- 6) Для подключения сторонних библиотек и фреймворков при разработке использовать сборщик проектов Maven или Gradle.

Список источников по темам, отсутствующим в лекциях

1. Про REST API:
 - <https://yandex.cloud/ru/docs/glossary/rest-api>
 - <https://gb.ru/blog/rest-api/>
2. Про формат JSON:
 - <https://tproger.ru/articles/что-такое-json-vvedenie>
3. Библиотека Jackson для работы с JSON:
 - <https://www.baeldung.com/jackson-object-mapper-tutorial>
4. HTTP клиент Apache HttpClient, который можно использовать для формирования запросов к конечным точкам:
 - <https://www.baeldung.com/apache-httpclient-cookbook>
 - <https://habr.com/ru/companies/umbrellaitcom/articles/423591/>
 - <https://mvnrepository.com/artifact/org.apache.httpcomponents.client5/httpclient5>
5. Списки открытых API:
 - <https://github.com/public-api-lists/public-api-lists>
 - <https://github.com/public-apis/public-apis>
 - <https://habr.com/ru/articles/769384/>
 - <https://developer.donnoval.ru/free-api/>