

## LU13.A03 Crawler Pipeline

Machen Sie einen Cherry-Pick des Commits TODO, um den fertigen WikiCrawlerPipelineService zu erhalten.

Nutzen Sie den folgenden Code, um die Funktionalität zu überprüfen und sämtliche vorhandenen Inhalte dieses Moduls in die Vektordatenbank zu speichern.

```
package ch.bzz.rag;

import ch.bzz.rag.service.*;
import lombok.extern.slf4j.Slf4j;
import org.springframework.boot.SpringApplication;
import org.springframework.boot.autoconfigure.SpringBootApplication;
import org.springframework.context.ConfigurableApplicationContext;

@Slf4j
@SpringBootApplication
public class ApplicationMain {

    public static void main(String[] args) {
        String namespaceUrl =
"https://wiki.bzz.ch/de/modul/ffit/3-jahr/java/";
        String regexFilter = "(?!.*\\s/start$).*";
        boolean overwriteExisting = false;

        ConfigurableApplicationContext ctx =
SpringApplication.run(ApplicationMain.class, args);
        WikiCrawlerPipelineService pipeline =
ctx.getBean(WikiCrawlerPipelineService.class);
        pipeline.runPipeline(namespaceUrl, regexFilter, overwriteExisting);
        ctx.close();
    }
}
```

From:

<https://wiki.bzz.ch/> - BZZ - Modulwiki

Permanent link:

<https://wiki.bzz.ch/de/modul/ffit/3-jahr/java/learningunits/lu13/aufgaben/a03?rev=1764636499>

Last update: 2025/12/02 01:48

