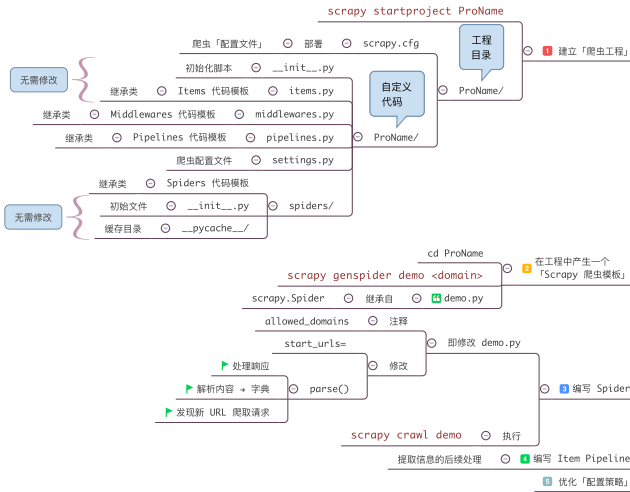
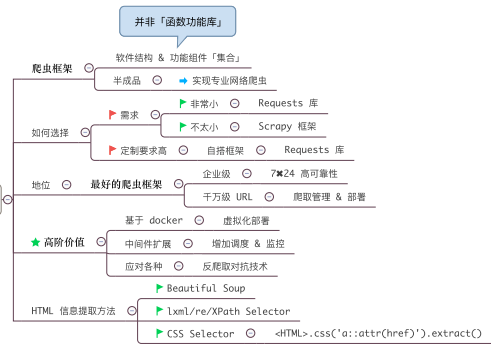


Requests	Scrapy	同! !
页面级 功能库 并发性不足 性能差 重点: 页面下载 定制灵活	网站级 框架 并发性好 性能高 重点: 爬虫结构 一般定制 OK, 深度定制困难	页面请求 & 爬取 同为两条重要技术路线 requests-bs4-re scrapy 处理 js 提交表单 应对验证码 PhantomJS

Requests vs Scrapy

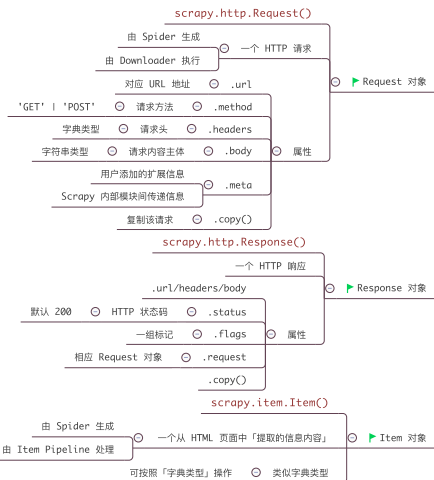
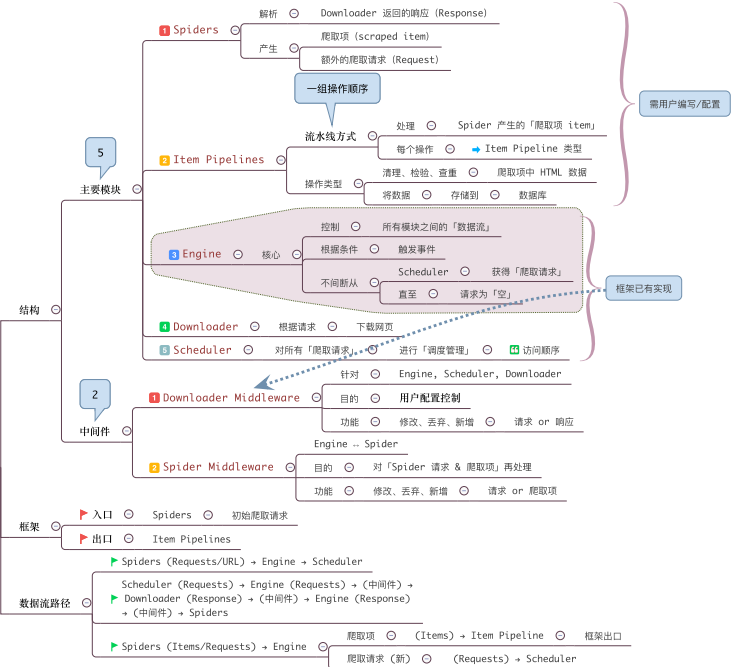
简介



使用步骤

公众号: 数林梵风
<https://woaieif.github.io/>
4 Scrapy 爬虫框架

“5+2” 结构



数据类型

常用命令

