Java爬取网站信息

主要api

httpclient httpresponse httpentity Pattern Matcher

1. 通过httpClient获取响应流
2. 将流转换为字符
3. 用Pattern制定正则规定
4. Matcher匹配结果