

Files



包含Java项目文件



Usage

该程序是一个简单的Web Crawler,爬取某个网站（如www.cs.zju.edu.cn）的网 页，可以指定**爬取的深度**，爬取后在命令行显示树状结构；通过命令行输入，可以对标题（title）和内容（content）进行内容检索，展示网页信息(**遍历后得到的所有url地址为最多200条，因为深度 ≥ 2 时，遍历后建立索引和doc时间过长，因此加以限制**)

Input root URL输入根地址

Input searching depth输入遍历深度

Input query string输入查询关键字

Input query field输入查询域