import org.jsoup.Jsoup;

import org.jsoup.nodes.Document;

import org.jsoup.nodes.Element;

import org.jsoup.select.Elements;

import java.io.BufferedReader;

import java.io.IOException;

import java.io.InputStreamReader;

import java.net.URL;

import java.net.URLConnection;

import java.util.ListIterator;

public class MySpider {

public static void main(String[] args) throws IOException {

URL url = new URL("http://www.lib.uestc.edu.cn/bysw/list60");

URLConnection urlConnection = url.openConnection();

BufferedReader bufferedReader = new BufferedReader(new InputStreamReader(urlConnection.getInputStream(), "utf-8"));

String line = null;

StringBuffer stringBuffer = new StringBuffer();

while ((line = bufferedReader.readLine()) != null) {

stringBuffer.append(line + "\n");

}

bufferedReader.close();

Document d = Jsoup.parse(stringBuffer.toString());

Elements divs = d.select("div.grid\_1\_of\_3.images\_1\_of\_3.top\_grid");

for (ListIterator<Element> it = divs.listIterator(); it.hasNext(); ) {

Element div = it.next();

Element h = div.getElementsByTag("h4").first();

String[] info = h.text().split("\\s+");

System.out.println("书名：" + info[0] + "，作者：" + info[1]);

}

}

}

