

7 爬虫 Http请求和Chrome

访问一个网页

<http://kaoshi.edu.sina.com.cn/college/scorelist?tab=batch&wl=1&local=2&batch=8&year=2013>

url : 协议 + 域名 / IP + 端口 + 路由 + 参数

ping

通过url能得到什么

在浏览器中打开

墙裂推荐大家使用Chrome浏览器

渲染效果、调试功能都是没话说的

<http://www.google.cn/intl/zh-CN/chrome/browser/desktop/index.html>

开发者工具

显示网页源代码、检查

1. Elements : 页面渲染之后的结构, 任意调整、即时显示;
2. Console : 打印调试;
3. Sources : 使用到的文件;
4. Network : 全部网络请求。

Http请求

Http是目前最通用的web传输协议

1. GET : 参数包含在url中;
2. POST : 参数包含在数据包中, url中不可见。

<http://shuju.wdzy.com/plat-info-59.html>

Url类型

1. html : 返回html结构页面, 通过浏览器渲染后呈现给用户;
2. API : Application Programming Interfaces, 请求后完成某些功能, 例如返回数据。

<http://kaoshi.edu.sina.com.cn/?p=college&s=api2015&a=getAllCollege>

我的个人网站 : <http://zhanghonglun.cn/>

雪晴数据网 : <http://www.xueqing.tv/course/73>

毕业僧 : <http://www.biyeseng.cn/front/couinfo/34>

研趣网 : <http://www.rinterest.cn/course/72>

优酷 : http://v.youku.com/v_show/id_XMTc1NDY1NjM2OA==.html?f=28356198&from=y1.7-3

腾讯 : <http://v.qq.com/vplus/334d685e3cea789f67d5bd9617482a49/foldervideos/1ad000101bqho2b>

土豆 : <http://www.tudou.com/listplay/1stkSpQCB1o.html>

爱奇艺 : http://www.iqiyi.com/v_19ruejtj7x.html?list=19rkraftq