网站爬虫的管理规则：授课老师：一叶

1. 链接长度（长的话爬虫就会少）：中文：（爬虫访问会有阻碍）

[www.一叶.com](http://www.一叶.com) www.一叶.com/老师

英文：[www.yiye.com](http://www.yiye.com)

1. 网站识别回访：打开速度快

识别难度 图片（很难）

视频（不识别）

----->文字多

1. 有效收录：有价值
2. 爬虫越多，越容易被收录（一个网站起码有20-30个链接）
3. 访问日志

rdseo9