示例

1. 进入政府（部门）官网

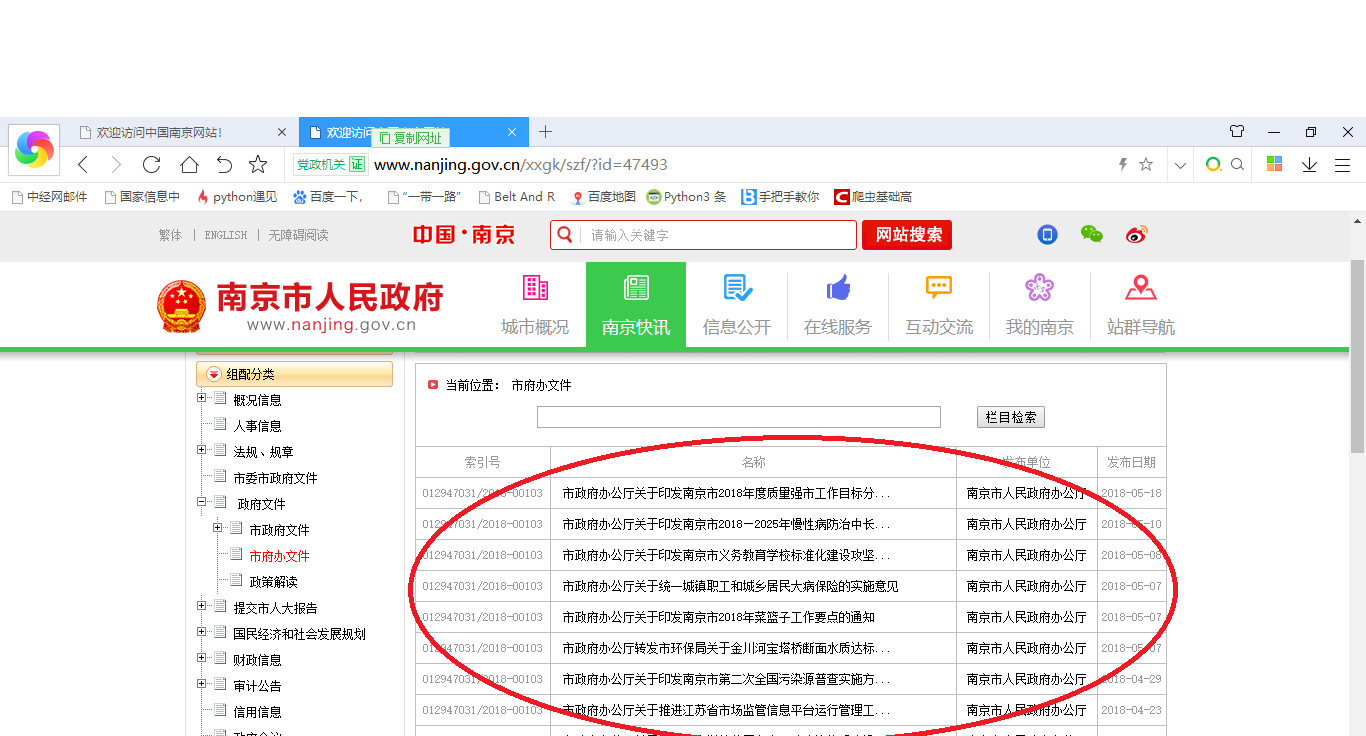
<http://www.nanjing.gov.cn/index.html>

找出其中的政府文件栏目，一般在信息公开中



1. 进入政府文件网站

<http://www.nanjing.gov.cn/xxgk/szf/?id=47493>



红色画圈部分为需要采集的文章。

采集范围为：市政府文件、市政府办公厅文件、部门文件、区文件

1. 进入文章页面



1. 将网页其中的部分内容爬取存储到本地TXT（名称为文件标题）

需采集网页URL、文件信息、文件内容

将网页中的文件下载存储到相应文件夹中，以文章标题命。

具体参考示例

网页中有下载链接或者附件的也需下载到相应的文件夹中

