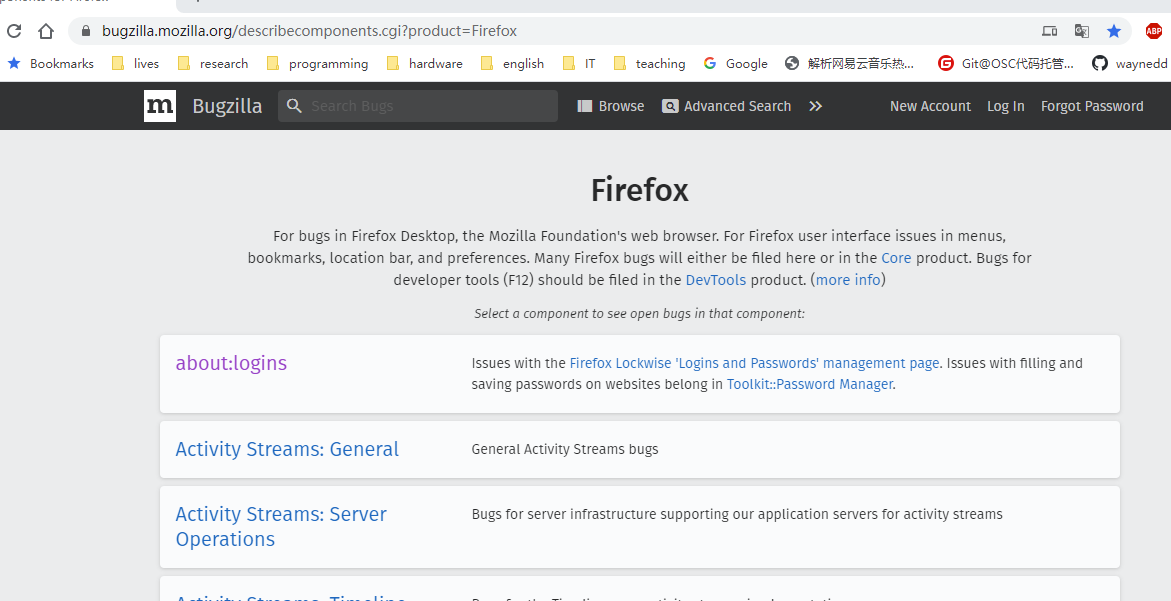
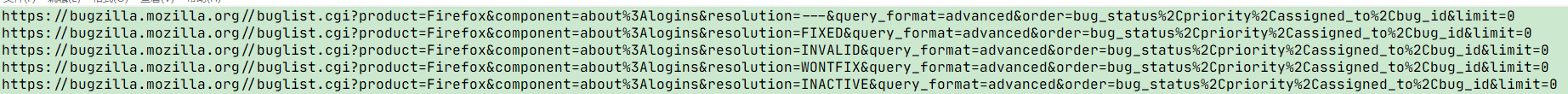
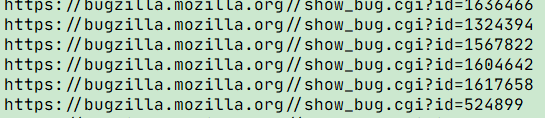
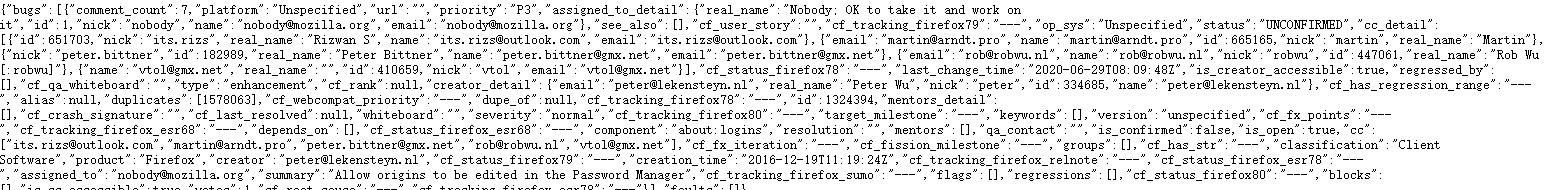
1. 首先用python获取<https://bugzilla.mozilla.org/describecomponents.cgi?product=Firefox>页面下的所有component的地址(SpiderBaseUrl.py)生成baseurl.txt文件:





1. 读取baseurl.txt文件的每一行获得相应component不同状态的所有基本bug信息的地址(spiderbaseinfosplit18\_19.py), 
2. 截取2中的bugid采用的方法是bugzilla提供的REST API接口读取相应的json格式的bug信息:https://bugzilla.mozilla.org/rest/bug/1324394



用python读取其中的有用字段存于csv文件中;

1. 同样用上述3的方法获取bug的history数据;
2. 对于user的profile由于buaziila没有提供REST API的接口,只能采用传统的python爬取数据的方法获取user的profile数据.