ESI学科数据获取及分析

实验目的

通过编程训练，能用一些爬虫工作实现网络数据的爬取。

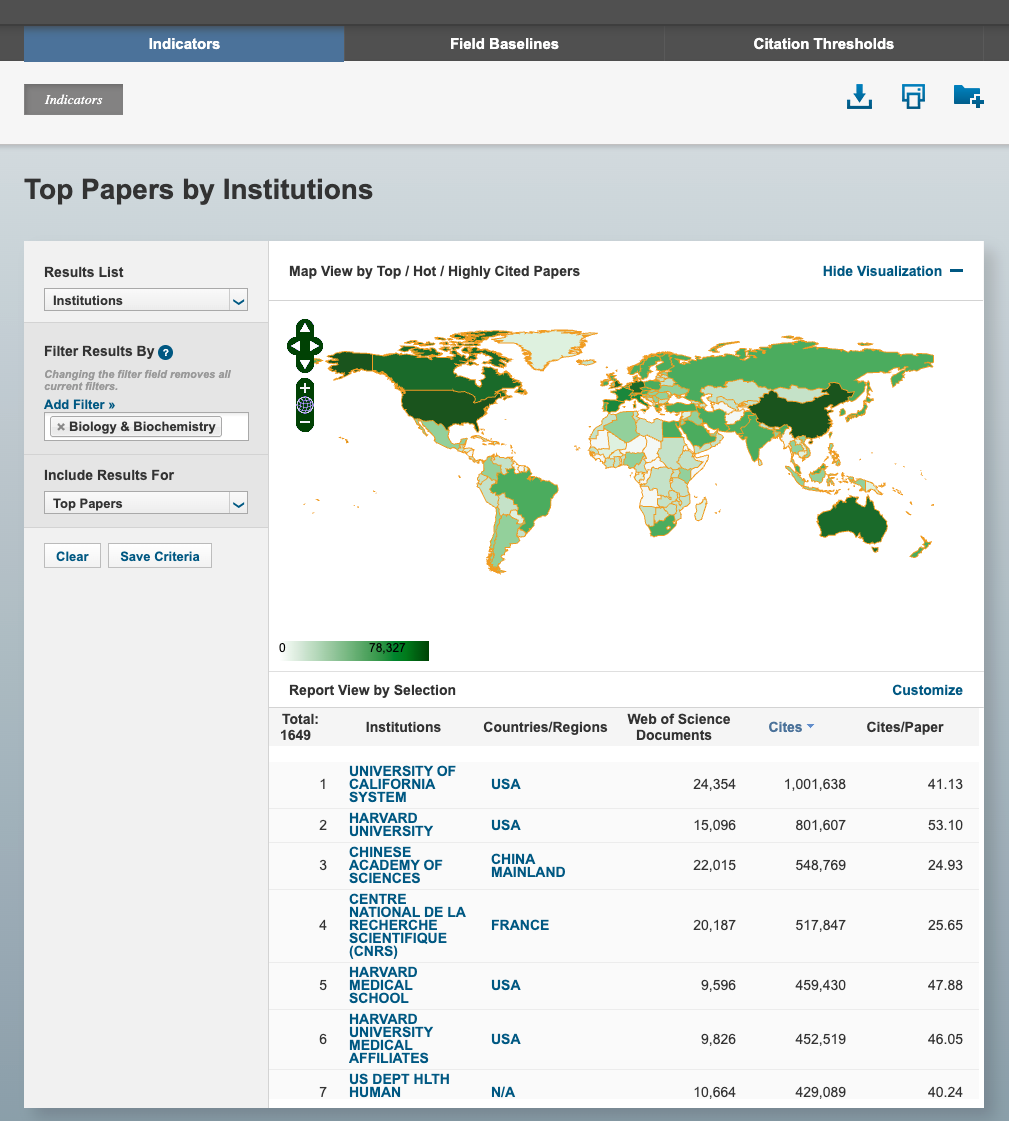
实验内容和要求：

使用Python语言，设计一个网络数据爬取工具，能够自动的获取ESI学科排名的相关数据。

ESI的网站如下：

<https://clarivate.com/academia-government/scientific-and-academic-research/research-funding-analytics/essential-science-indicators/>

通过CERNET CHINA，使用华师大个人账号登录，会进入如下界面。



1. 请写一个爬虫程序，能够自动爬取页面内的数据。例如，Results List下拉选择Institutions, 通过自定义Filter Results By中的Research Fields, 实现对不同学科的全球大学排行都分别下载下来。
2. 通过爬虫获取的数据，对华东师范大学的学科情况进行数据分析，并形成一份报告文件。
3. 将获取的数据导入到一个关系型数据库系统中（系统可以自选）
4. 优化关系型数据，并整理一个合理的schema
5. 通过写SQL语句，获取华东师范大学在各个学科中的排名
6. 通过写SQL语句，获取中国（大陆地区）大学在各个学科中的表现
7. 通过写SQL语句，分析全球不同区域在各个学科中的表现