**大规模爬虫平台**

**主要模块：下载模块 析取模块 存储模块**

**下载模块：**提供一个易用且强大的校园网络爬虫,让网页数据采集变得前所未有的简单.可以自动采集校园信息服务的页面内容。

**析取模块:**对页面进行分析，关键字提取过滤和分割，表格的分析提取等。解析后以便程序做下一步的处理。

**存储模块：**通过对数据的分析，建设一个丰富且新鲜的数据资源库.