

易律——法律信息检索工具

一、成员介绍

金超轶 电子科学与技术 2016 级
毛子健 计算机科学与技术 2016 级
罗宇彬 通信工程 2016 级
赵宇清 物理学 2014 级
陈甸甸 法学 2017 级
蒲星宇 无机非金属材料工程 2017 级

二、作品简介

易律是我们制作的一款提供搜索法律信息功能的检索工具,大家可以通过我们的软件去搜索自己想要了解的法律相关知识,并且每个法律条文都会为您提供相应的法律案例,以助你更深入地理解。

易律的内容会不定期进行更新,确保用户获得正确和最新的法律信息。

易律还提供论坛和留言功能,想要参与讨论功能用户需要先进行登陆操作。

三、应用场景

针对于学生以及对法律知识不熟悉的人群,需要学习和查询相关法律信息知识案例时使用。

四、总体设计

a)基本思路

将爬虫技术,网页构建和数据库技术三者生态融合

b)主要技术和解决方案

Flask 网页框架, generate_password_hash 加密技术, context_processor 钩子函数,
pymysql(连接数据库)

其他技术: 数据库 (mysql), 网页制作 (h5,css, js)

五、特色和创新点

a)特色

对我们日常生活有很大帮助

成本低,操作简单

以行为为关键词的检索系统

以网站形式无人自动化查询,快捷零成本

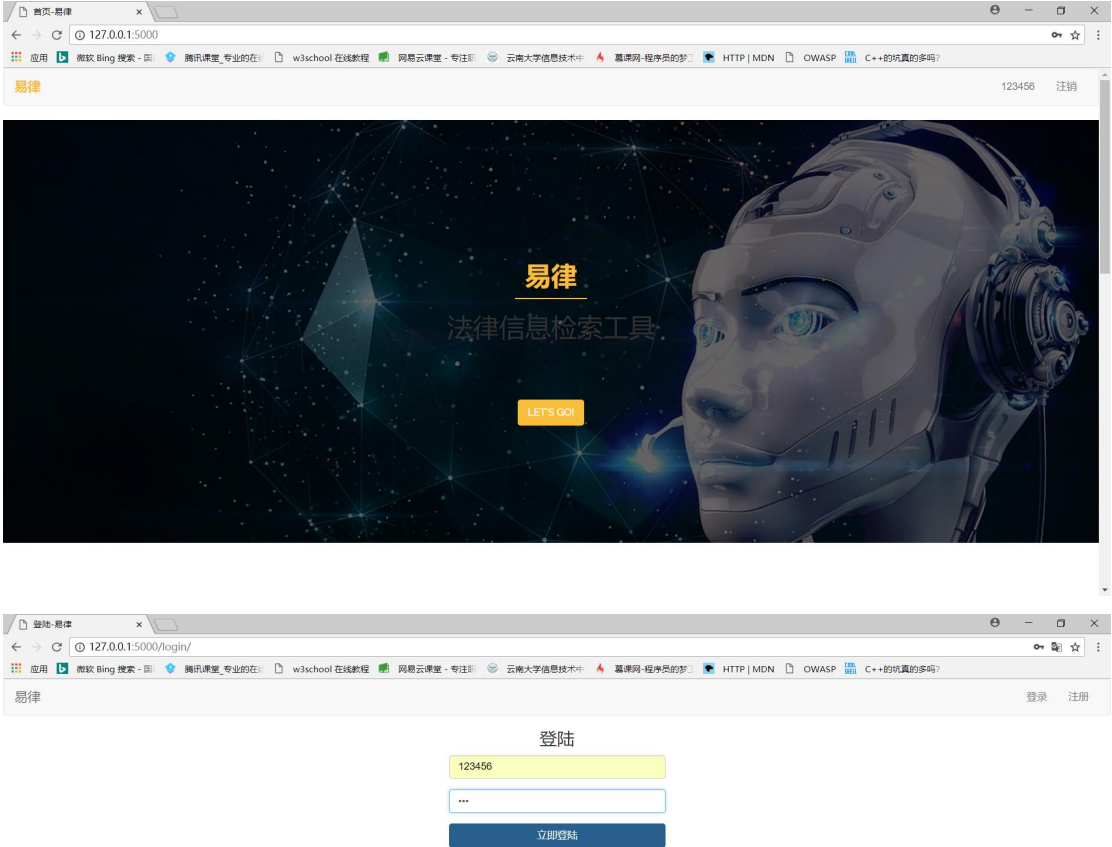
b)创新

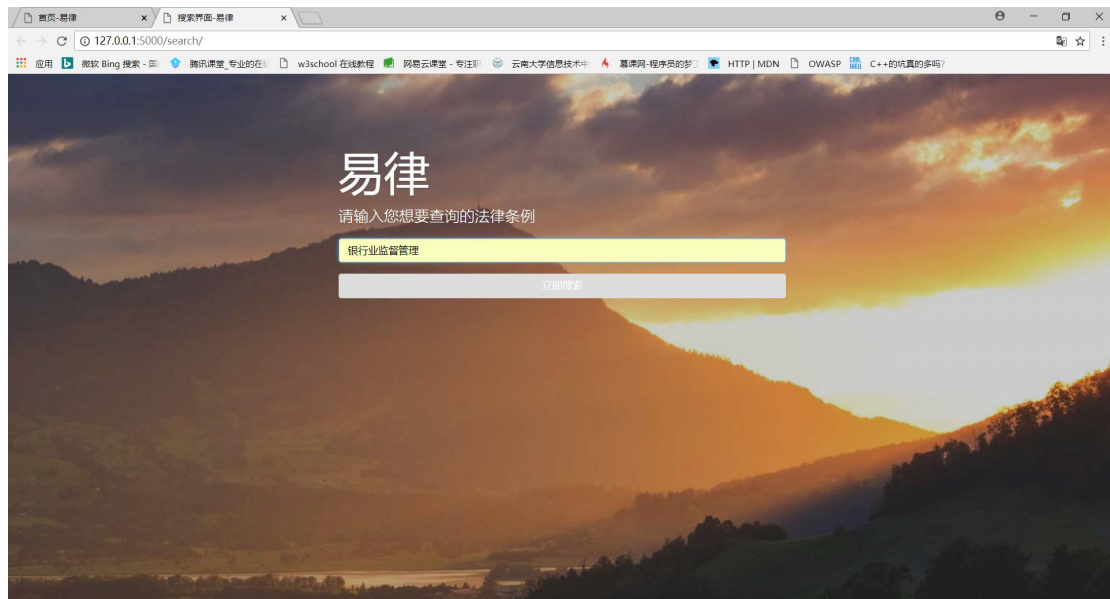
利用爬虫技术内容可以不断实现更新

提供登陆,注册,留言等功能

利用了钩子函数等其他自学的内容，与 python 结合
以 python 爬虫+数据库技术涵盖海量法律及行政法规

六、运行截图





七、项目组成员的工作心得

成员一：金超轶，2016 级，项目负责人

- 学习能力方面：能有效找寻适当的学习方法，去询问老师从哪些方面着手这个作品，在理清设计思路和过程后，能积极学习相关知识。查找相关书籍，百度搜索资料，并自学网络上的课程进行补充。能够和小组成员积极讨论，寻求解决方法。
- 独立思考能力方面：能独立思考问题，运用能够查询的资料来解决问题。能思考代码运行的逻辑，和整个作品的运行框架，能忍受长时间代码错误并寻求解决方案。
- 发现问题能力方面：能仔细试验每个功能的代码块，寻找可能出错的地方。如果存在错误就积极修正并尝试解决。
- 其他方面：整个过程还是存在一些不足，虽然进行了合理和明确的分工，但没能把全部功能实现（还有爬虫更新数据库没能实现）稍微有点遗憾。毕竟大家不是一个班的同学，沟通交流还是比较麻烦的，而且这个课不是唯一的事情要做，很多知识都是从头学

起,所以存在不足也是情有可原的。但是我们组还是组织了几次在图书馆的沟通,我也和负责技术的同学进行了沟通,再加上自己的努力,算是比较圆满的实现了作品的功能。这次小组合作加深了我对合作的理解,让我认识到了自己的不足,也学习了很多新的知识,如 flask, MySQL, h5 语言等等,虽然过程比较辛苦,但收获也很多。

成员二: 罗宇彬,2016 级,负责数据库设计制作

- a)学习能力方面:通过本次项目,我加深了对以前学过的数据库设计的理解,对 SQL 语言有了更深层次的认识,对 python 结合 SQL 开发一些项目有了一些大体的认识,相信以后再制作相同的或类似的项目会更加得心应手。
- b)独立思考方面:在数据库设计的过程中,我先对项目负责人和前端开发的同學制作的前端网页功能做了一个功课,弄明白了想如何实现项目中需要的功能有哪些,然后对数据库的表和列以及需要的内容进行了添加,设计了法律,用户,管理员,留言板等几个表。
- c)发现问题能力方面:在创建好数据库之后,代码出现了许多的问题,例如:标点符号不兼容,SQL sever 和 MYSQL 语法不兼容,行列未对应的这些问题,我都一一解决了,最后整合的时候,我的版本和负责人的版本不一样,导致不兼容,最后,多次修改之后,得以顺利运行。
- d)其他方面:通过本次项目,我发现,只要大家都努力对项目成功做出贡献,项目的结果不会差,所以,不管是小组还是个人,任务必须要努力去完成。

成员三: 赵宇清, 2014 级, 爬虫部分

- a)学习能力方面: 学习了爬虫如何实现, 但遗憾的是没能最终实现
- b)独立思考能力方面: 能独立学习和思考爬虫技术
- c)发现问题能力方面: 能发现爬虫技术的难点
- d)其他方面: 在小组讨论中能积极讨论

成员四: 毛子健, 2016 级, 网页设计负责人

- a)学习能力方面: 学会了使用 python 连接数据库, 熟练运用 python 的 flask 模块实现各种功能。
- b)独立思考能力方面: 为了设计更好的网页去参考了 bootstrap 框架并学习了少部分 Javascript。
- c)发现问题能力方面: 没有太多的感想, 主要由组长统一大家的问题。
- d)其他方面: 团队协作是一件很重要的事。

成员五: 陈甸甸, 2017 级, PPT 部分

主要负责产品设计和 PPT 制作及演讲工作。在讨论初期提出制作爬虫爬取法律条文的构思,不断完善产品定位,提出日常性法律条文检索的定位,提议关键词检索的设计。中期设计展示内容,将部分法律条文加入数据库以提升展示的效果。后期学习产品技术逻辑,选取视频,收集各人员工作分配,制作出 PPT, PPT 设计为突出产品背景和定位、重点讲解亮点 python 技术并辅以产品展示和团队介绍的结构。最后在课堂上进行 PPT 演讲,和参与产品问答解释。

成员六: 蒲星宇, 2017 级, 文案部分

- a)学习能力方面: 对自身要求还不是很严格, 还需要多培养自学能力
- b)独立思考能力方面: 能够提出一些自己的见解, 对问题有自己的看法
- c)发现问题能力方面: 还不是特别仔细耐心, 还需继续努力

八、存在的问题

对爬虫技术还是不是特别熟悉，部分需要用爬虫实现的功能还未完成

九、代码

a)其他部分：单独附文件

b)爬虫部分，但还不能更新数据库

```
# -*- coding: utf-8 -*-
```

```
import urllib.request
import xmltodict
import re
from bs4 import BeautifulSoup
```

```
rootUrl = "[图片]http://www.chinalaw.gov.cn"
html=urllib.request.urlopen("[图
片 ]http://www.chinalaw.gov.cn/module/web/jpage/dataproxy.jsp?startrecord=0&endrecord
=100000&perpage=100000&col=1&appid=1&webid=1&path=%2F&columnid=11&source
ContentType=1&unitid=1648&webname=%E4%B8%AD%E5%9B%BD%E6%94%BF%E5%BA%9C%E6
%B3%95%E5%88%B6%E4%BF%A1%E6%81%AF%E7%BD%91&permissiontype=0")
doc = xmltodict.parse(html)
pattern=r'.*href="(.*)">(.*?)<.*'
match=re.match(pattern, doc['datastore']['recordset']['record'][0])
url = "{rootUrl}{path}".format(rootUrl=rootUrl,path=match.group(1))
print (url)
print (match.group(2))
html=urllib.request.urlopen(url)
soup = BeautifulSoup(html)
law = soup.find("div", {"id":"zoom"})
print("-----")
#print (type(law))
print (u"".join(law.findAll(text=True)))
print(law.text)
print("-----")
```