

Python语言程序设计

实例12: 政府工作报告词云



嵩 天 北京理工大学





问题分析

直观理解政策文件

- 需求: 对于政府工作报告等政策文件, 如何直观理解?

-体会直观的价值: 生成词云 & 优化词云

政府工作报告等文件 🖒 有效展示的词云

问题分析

《决胜全面建成小康社会 夺取新时代中国特色社会主义伟大胜利》

在中国共产党第十九次全国代表大会上的报告

(2017年10月18日)

习近平

https://python123.io/resources/pye/新时代中国特色社会主义.txt

问题分析

《中共中央 国务院关于实施乡村振兴战略的意见》

2018—号文件

(2018年01月02日)

中共中央 国务院

https://python123.io/resources/pye/关于实施乡村振兴战略的意见.txt



政府工作报告词云

基本思路

- 步骤1: 读取文件、分词整理

- 步骤2: 设置并输出词云

- 步骤3:观察结果,优化迭代

```
#GovRptWordCloudv1.py
import jieba
import wordcloud
f = open("新时代中国特色社会主义.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = " ".join(ls)
w = wordcloud.WordCloud( font path = "msyh.ttc",\
   width = 1000, height = 700, background color = "white", \
w.generate(txt)
w.to file("grwordcloud.png")
```



新时代中国特色社会主义

```
#GovRptWordCloudv1.py
import jieba
import wordcloud
f = open("关于实施乡村振兴战略的意见.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = " ".join(ls)
w = wordcloud.WordCloud( font path = "msyh.ttc",\
   width = 1000, height = 700, background color = "white", \
w.generate(txt)
w.to file("grwordcloud.png")
```



2018年一号文件

```
#GovRptWordCloudv1.py
import jieba
import wordcloud
f = open("新时代中国特色社会主义.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = " ".join(ls)
w = wordcloud.WordCloud( font path = "msyh.ttc",\
   width = 1000, height = 700, background color = "white", \
    max words = 15)
w.generate(txt)
w.to file("grwordcloud.png")
```



新时代中国特色社会主义

```
#GovRptWordCloudv1.py
import jieba
import wordcloud
f = open("关于实施乡村振兴战略的意见.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = " ".join(ls)
w = wordcloud.WordCloud( font path = "msyh.ttc",\
   width = 1000, height = 700, background color = "white", \
   max words = 15)
w.generate(txt)
w.to file("grwordcloud.png")
```



2018年一号文件

准备好电脑,与老师一起编码吧!



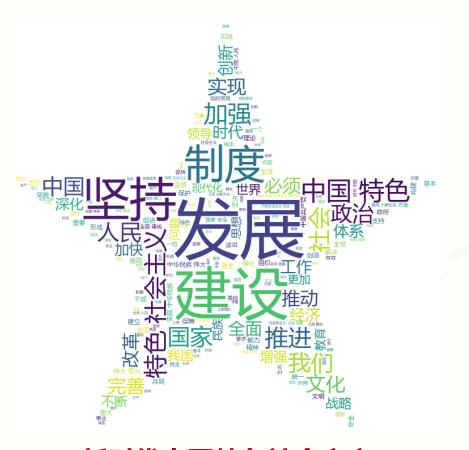
政府工作报告词云

更有形的词云





```
#GovRptWordCloudv2.py
import jieba
import wordcloud
from scipy.misc import imread
mask = imread("fivestart.png")
f = open("新时代中国特色社会主义.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = "".join(ls)
                          font_path = "msyh.ttc", mask = mask\
w = wordcloud.WordCloud(
   width = 1000, height = 700, background color = "white", \
w.generate(txt)
w.to file("grwordcloud.png")
```



新时代中国特色社会主义



2018年一号文件

```
#GovRptWordCloudv2.py
import jieba
import wordcloud
from scipy.misc import imread
mask = imread("chinamap.jpg")
f = open("新时代中国特色社会主义.txt", "r", encoding="utf-8")
t = f.read()
f.close()
ls = jieba.lcut(t)
txt = "".join(ls)
                          font_path = "msyh.ttc", mask = mask\
w = wordcloud.WordCloud(
   width = 1000, height = 700, background color = "white", \
w.generate(txt)
w.to file("grwordcloud.png")
```





新时代中国特色社会主义

2018年一号文件



举一反三

扩展能力

- 了解wordcloud更多参数,扩展词云能力

- 特色词云: 设计一款属于自己的特色词云风格

- 更多文件: 用更多文件练习词云生成

