



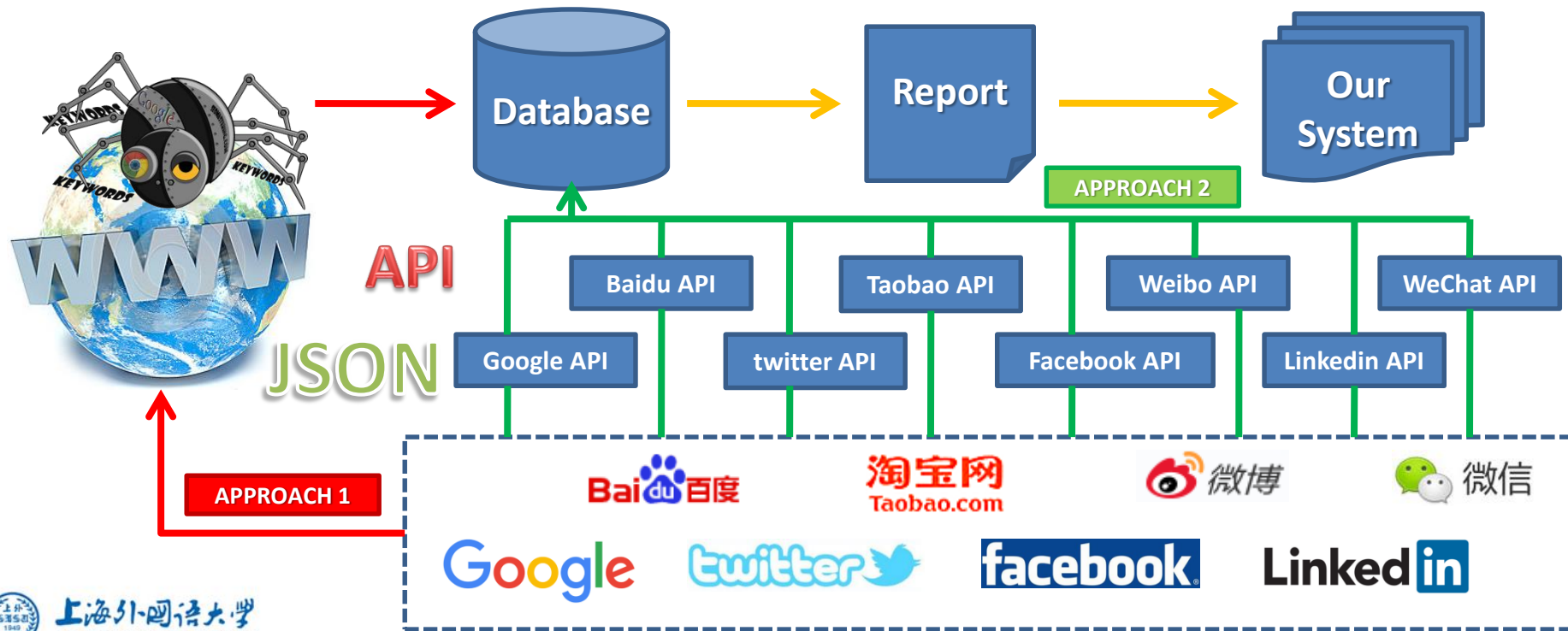
New Media Data Analytics and Application

Lecture 7: Data Collection via Web API

Ting Wang

Objectives

How to get data from social media?



- Web API
- JSON
- Examples of Web API Applications





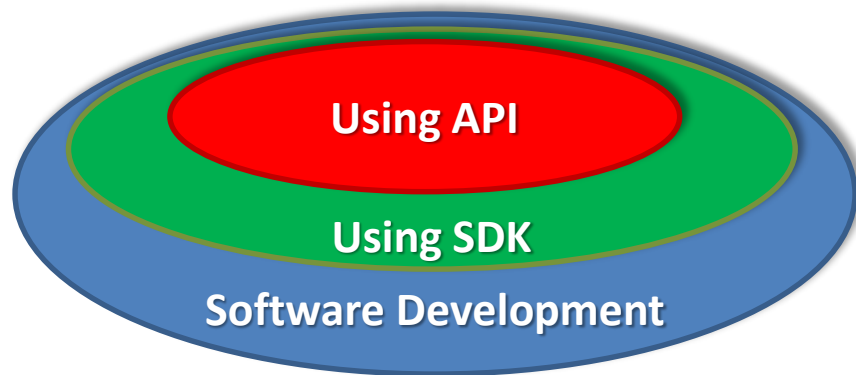
an access to the web site data

Web API

SDK (Software Development Kit)

软件开发工具包

a set of
software development tools,
source code,
instructions,
examples,
and ...



How to print a booklet for our sport games?



具体的

2

Using SDK
在人家的套路基础上走

1



Traditional Software
Development
自己设计套路走

抽象的

3

Using Web API
自己不动脑子，
完全按人家的套路走

淘宝网 Taobao.com

宝贝 彩页设计印刷 搜索

在结果中排除 请输入要排除的词 确定

品牌印刷	益好	五折促销	东讯
免设计费 三星冠厂家 全国包邮 500张 — 25元 1000张 — 30元 2000张 — 80元 3000张 — 90元 5000张 — 150元 画册印刷 买贵退差价 ¥68.00起 1350人付款 宣传册印刷彩页印刷企业宣传册画册设计制作A4单页折页广告彩印 三益旗舰店 上海	足克足量 定稿后次日发货 手机下单 500张 45元 宣传册/彩页/海报/名片印刷 免费设计 A5单页1千张58元 ¥68.00起 7583人付款 宣传册印刷彩页印刷画册印刷传单免费设计制作广告打印三折页A4单页 三益旗舰店 广东 广州	五折促销 A4宣传单/海报印刷 免费设计 500张 — 45元 1000张 — 70元 2000张 — 148元 3000张 — 218元 5000张 — 298元 买贵退差价 足克足量 ¥68.00起 5745人付款 A4宣传册印刷彩页印刷500单页广告彩印传单设计制作画册海报三折页 东讯旗舰店 江苏 南京	源拓印刷 A4宣传单印刷 免费设计 足克足量 500张 40元 1000张 — 70元 2000张 — 148元 买贵退差价 次日发货 (自单) ¥68.00起 5249人付款 宣传册印刷传单画册印刷彩页广告设计制作折页海报A4单页宣传册 源拓旗舰店 广东 深圳

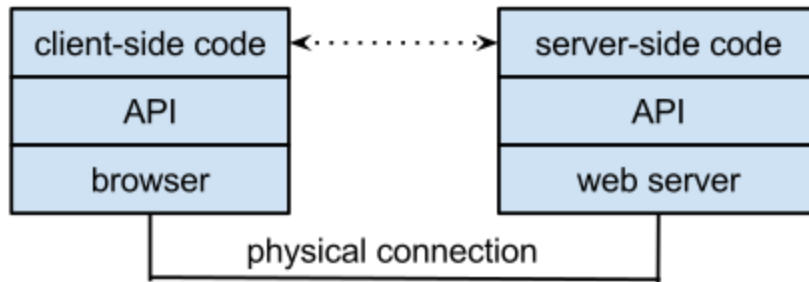
API (Application Programming Interface)

应用程序接口

a set of
subroutine definitions,
protocols,
and tools

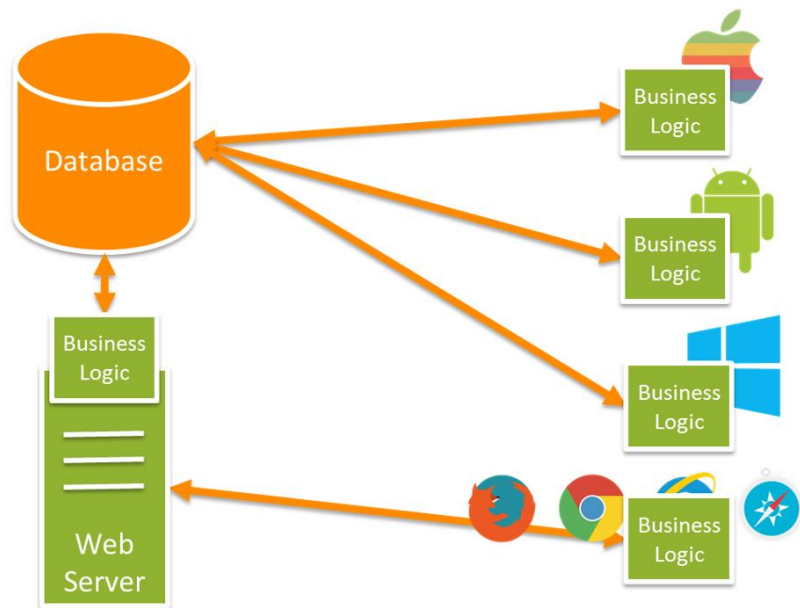
- What is Web API?

BS/CS

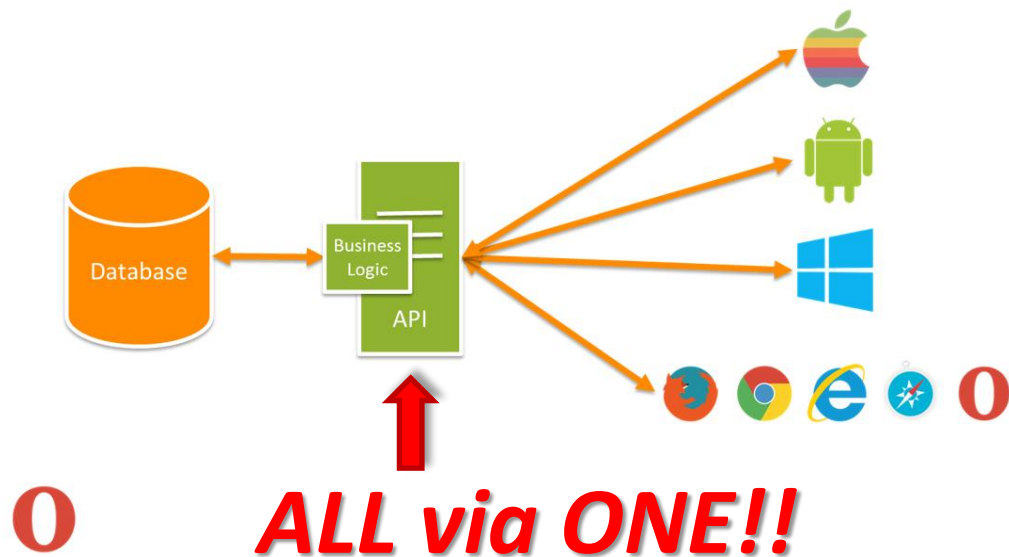


Web API

Short Short Ago



Now



Integrating all the operations into one package,
including “insert”, “delete”, “update”, and “select”.

In Unified Style, two crucial elements:

– Unified Method :

- HTTP (HyperText Transfer Protocol)

超文本传输协议

– Unified Language :

- JSON
- XML



HTTP(HyperText Transfer Protocol)

超文本传输协议

an application protocol for distributed, collaborative, hypermedia information systems, which is the foundation of data communication for the World Wide Web.

- GET
 - Give Me Information
- POST
 - Give Server Information

Do you remember
“Login.py” ?



Web API

```
1 #!/usr/bin/env python3
2 # -*- coding: utf-8 -*-
3
4 from flask import Flask
5 from flask import request
6 import pymysql
7 import pymssql
8 import sqlite3
9
10 app = Flask(__name__)
11
12 @app.route('/', methods=['GET', 'POST'])
13 def home():
14     return '''<h1>Home</h1>
15             <p><a href="/login">Login</a></p>'''
16
17 @app.route('/login', methods=['GET'])
18 def login_form():
19     return '''<form action="/login" method="POST">
20             <p><input name="username"></p>
21             <p><input name="password" type="password"></p>
22             <p><button type="submit">Log In Now</button></p>
23             </form>'''
24
25
26 @app.route('/login', methods=['POST'])
27 def login():
28     # get data from database:
29     # 1.Connection Open
30     conn = pymysql.connect(user='sa', password='123456', database='login')
31     # 2.Cursor Creating:
32     cursor = conn.cursor()
33     # 3.SQL Execution
34     sqlstr = 'select * from customer where customer_name=\''+request.form['username']+\'
35             |\' and customer_pwd=\''+request.form['password']+\'\'
36     # 4.Cursor Moving
37     cursor.execute(sqlstr)
38     numrows = len(cursor.fetchall())
39     conn.commit()
40     # 5.Connection Close
41     cursor.close()
42     conn.close()
43     # get data from request:
44     if numrows>0:
45         return '<h3>Hello, '+request.form['username']+ '</h3>'
46     return '<h3>Bad username or password.</h3>'
47
48 if __name__ == '__main__':
49     app.run()
```

Language of Web API

– XML (EXtensible Markup Language)

```
<?xml version="1.0" encoding="UTF-8" ?>
  <employees>
    <firstName>Bill</firstName>
    <lastName>Gates</lastName>
  </employees>
  <employees>
    <firstName>George</firstName>
    <lastName>Bush</lastName>
  </employees>
  <employees>
    <firstName>Thomas</firstName>
    <lastName>Carter</lastName>
  </employees>
```



Ref. <http://www.w3school.com.cn/x.asp>

Language of Web API

– JSON (JavaScript Object Notation)

```
{ "employees": [  
  { "firstName":"Bill" , "lastName":"Gates" },  
  { "firstName":"George" , "lastName":"Bush" },  
  { "firstName":"Thomas" , "lastName":"Carter" }  
]}
```

Ref. <http://www.w3school.com.cn/json/index.asp>



Transfer between XML and JSON

<http://www.bejson.com/xml2json/>

在线XML、JSON数据互转

如题

XML

```
<?xml version="1.0" encoding="UTF-8" ?>
<employees>
  <firstName>Bill</firstName>
  <lastName>Gates</lastName>
</employees>
<employees>
  <firstName>George</firstName>
  <lastName>Bush</lastName>
</employees>
<employees>
  <firstName>Thomas</firstName>
  <lastName>Carter</lastName>
</employees>
```

JSON

```
{ "employees": [ { "firstName": "Bill", "lastName": "Gates" }, {
"firstName": "George", "lastName": "Bush" }, { "firstName": "Thomas",
"lastName": "Carter" } ] }
```

转义编码



清空





Ask A Question

Why JSON is more and more popular?



EXAMPLE 1: News 1

I ♥ APIs

Get this free sticker at api.gee.com

Web API

www.juhe.cn



Juhe is a technology company that offers mobile data services and products in China.

API分类

全部 生活常用 车辆服务 金融征信 位置服务 即时通讯 应用开发 旅游出行 站长工具 文娱体育 充值缴费 礼品卡

三 全部

输入想要检索的API



排序: 综合 最新 热门 付费类型: 全部 免费 收费

 短信API服务 ¥0.0301/次	 银行卡四元重校验 ¥0.1500/次	 全国车辆违章 ¥0.0400/次	 汽车维修保养记录 ¥6.0000/次	 银行卡类型及真伪查询 ¥0.0070/次
 三网手机实名认证	 身份证实名认证	 银行卡三元重检测	 银行卡二元重检测	 违章代缴v2.0

联系·咨询

微信

京东

SUNING 苏宁

中国平安
PINGAN

万达集团

360
www.360.cn

同程旅游

滴滴
DIDI

dji 大疆

科大讯飞
iFLYTEK

中国大地保险
China Continent Insurance

蔚来

汽车之家
autohome.com.cn

融360
Rong360.com

HNA 海航集团

华大基因
BGI



上海外国语大学
SHANGHAI INTERNATIONAL STUDIES UNIVERSITY

Step 1: you should have your own account

- Log in the web:
<https://www.juhe.cn/register>
- You may register your account by cellphone or email.
- You may log in the website after you register it.

欢迎注册

百款接口免费试用

☒ 我已阅读并接受《聚合用户服务协议》

[通过邮箱注册>>](#)



Step 2: select an API(1)

- You may select a FREE API for test.



Step 2: select an API(2)

- Do not forget to confirm your information in two months, if the API is not free.



当前选择的数据为: 生活常用 > 彩票开奖结果查询

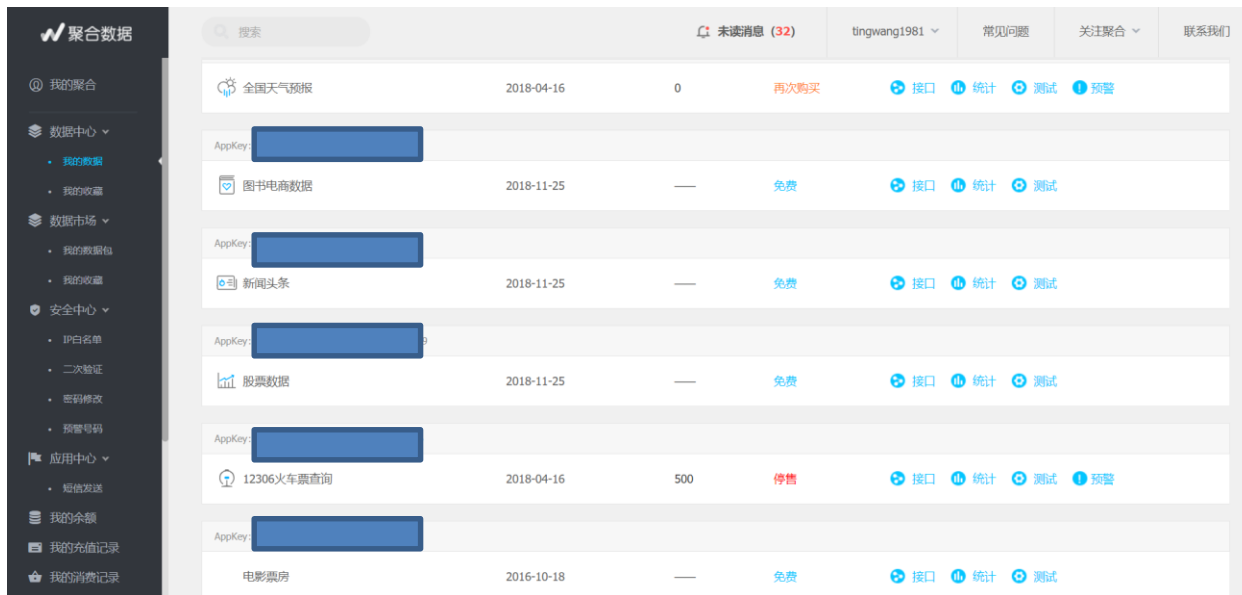
初始赠送: 免费使用 (根据会员等级不同, 每日可调用次数不同)

立即申请



Step 3: Save the AppKey of API

- You may start to use your API after you confirm your information.



The screenshot shows the '聚合数据' (Aggregated Data) website dashboard. The left sidebar contains navigation links: '我的聚合', '数据中心', '数据市场', '安全中心', and '应用中心'. The main content area displays a table of APIs with columns for API name, date, status, and actions. The AppKey for each API is highlighted with a blue box.

API名称	日期	状态	操作
全国天气预报	2018-04-16	0	再次购买
AppKey: [redacted]			
图书电商数据	2018-11-25	—	免费
AppKey: [redacted]			
新闻头条	2018-11-25	—	免费
AppKey: [redacted]			
股票数据	2018-11-25	—	免费
AppKey: [redacted]			
12306火车票查询	2018-04-16	500	停售
AppKey: [redacted]			
电影票房	2016-10-18	—	免费
AppKey: [redacted]			

- You may visit the main page of your API to get more detailed information.
- If the API is not free, you should pay for it.



全国天气预报 聚合数据 支持HTTPS

全国3000多个省市的实时天气预报, 未来7天、常用生活指数等

1000次 50000次 5000000次 300000次

¥1.00 (约0.0010元/次), 每人最多限购1份:

续费购买 收藏

产品功能 API文档 错误码参照 示例代码 联系我们

功能介绍 / Introduction

- 当前实况天气, 温度、湿度、风向风力
- 生活指数, 穿衣指数、洗车指数、晨练指数等
- 支持GPS坐标直接查询天气
- 覆盖全国2567市县区

全国天气预报API为企业提供准确、稳定、丰富的天气数据服务。通过API接口, 天气数据可以被轻松整合进手机应用、桌面应用或网站。聚合数据API整合多个数据源, 智能融合天气数据, 让数据更加准确, 服务更加稳定。

Step 4: write the code

- Examples are important.
- Sometimes, there is no Python examples.

How to write python code,
if we cannot find it in examples?
To look for a similar one!

全国天气预报 聚合数据 支持HTTPS

全国3000多个省市的实时天气预报, 未来7天、常用生活指数等

1000次 50000次 500000次 300000次

¥1.00 (约0.0010元/次), 每人最多限购1份:

续费购买 收藏

产品功能 API文档 错误码参照 示例代码 联系我们

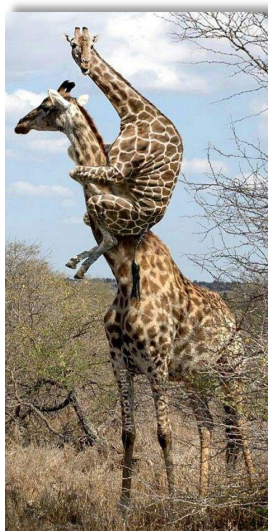
语言	标题	提供者	时间
PHP	全国天气预报API调用示例	JoneDwyane	2015-05-26 15:46:47
C#	天气接口调用示例	JoneDwyane	2015-09-15 14:22
Go	天气接口调用示例	JoneDwyane	2015-09-15 17:43
JAVA	天气接口调用示例	JoneDwyane	2015-09-26 18:16

常见问题:

内容	详细
常见问题:	http://www.juhe.cn/qa/index/1523867629

Web API

- Here is an example which has a code in Python.
<https://www.juhe.cn/docs/api/id/11>



Standing on the Shoulders of Giants
(SSG)

不要什么都去自己写，
要thinking in SSG!



完整教学代码示例:

语言	标题	提供者	时间
PHP	手机号码归属地查询	JoneDwyane	2015-09-01 15:51:54
Python	手机号码归属地接口调用示例	JoneDwyane	2015-09-14 19:35:17
C#	手机号码归属地接口调用示例	JoneDwyane	2015-09-15 13:34:19
Go	手机号码归属地接口调用示例	JoneDwyane	2015-09-15 19:17:39
JAVA	手机号码归属地接口调用示例	silik	2016-02-26 10:08:46



Web API

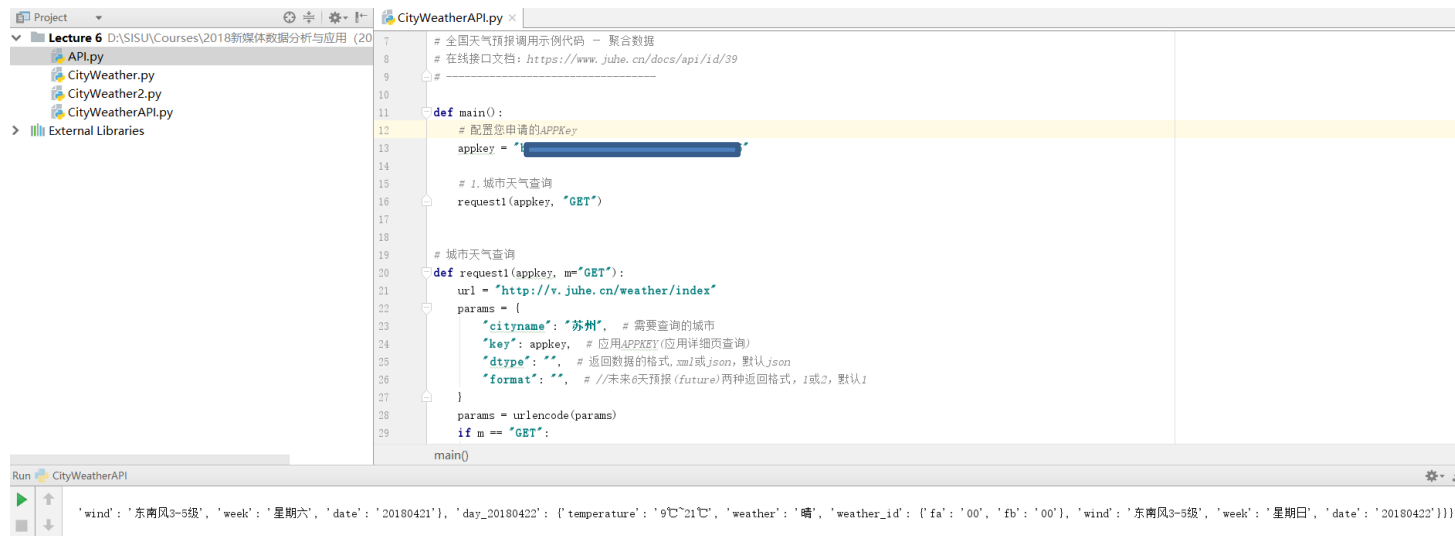
- Thinking in SSG!

- ① Analyze Code: <https://code.juhe.cn/docs/1018>

- ② Rewrite Code: based on both python examples from other APIs and the codes in your API.

- ③ Test Code

Some Codes
are in Python 2,
which should
be rewritten
with Python 3.



```
7 # 全国天气预报调用示例代码 - 聚合数据
8 # 在线接口文档: https://www.juhe.cn/docs/api/id/39
9 #
10
11 def main():
12     # 配置您申请的APPKey
13     appkey = "..."
14
15     # 1. 城市天气查询
16     request1(appkey, "GET")
17
18
19 # 城市天气查询
20 def request1(appkey, m="GET"):
21     url = "http://v.juhe.cn/weather/index"
22     params = {
23         "cityname": "苏州", # 需要查询的城市
24         "key": appkey, # 应用APPKEY (应用详细页查询)
25         "dtype": "", # 返回数据的格式, xml或json, 默认json
26         "format": "", # //未来6天预报 (future) 两种返回格式, 1或2, 默认1
27     }
28     params = urlencode(params)
29     if m == "GET":
30         main()
31
32
33 Run CityWeatherAPI
34
35 'wind': '东南风3-5级', 'week': '星期六', 'date': '20180421', 'day_20180422': {'temperature': '9℃~21℃', 'weather': '晴', 'weather_id': ['fa': '00', 'fb': '00'], 'wind': '东南风3-5级', 'week': '星期日', 'date': '20180422']})
```



- The Code

```
# -*- coding: utf-8 -*-
import json, urllib.request
from urllib.parse import urlencode

# -----

# 新闻头条调用示例代码 — 聚合数据
# by Dr. Ting Wang
# 在线接口文档: https://www.juhe.cn/docs/api/id/235
# -----

def main():
    # 配置您申请的APPKey
    appkey = "04f*****17c"

    # 1. 新闻头条查询
    request1(appkey, "GET")
```

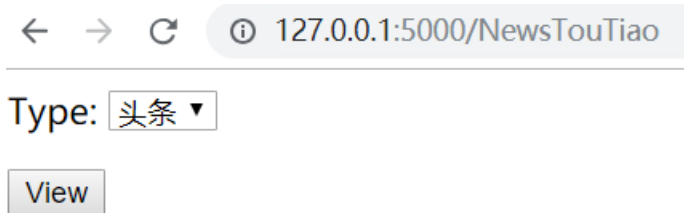
```
def request1(appkey, m="GET"):
    url = "http://v.juhe.cn/toutiao/index"
    params = {
        "type": "shehui", # top(头条, 默认), shehui(社会), guonei(国内), guoji(国际), yule(娱乐), tiyu(体育), junshi(军事), keji(科技), caijing(财经), shishang(时尚)
        "key": appkey, # 应用APPKEY(应用详细页查询)
        "dtype": "", # 返回数据的格式, xml或json, 默认json
    }
    params = urlencode(params)
    if m == "GET":
        f = urllib.request.urlopen("%s%s" % (url, params))
    else:
        f = urllib.request.urlopen(url, params)
    content = f.read()
    res = json.loads(content)
    if res:
        error_code = res["error_code"]
        if error_code == 0:
            # 成功请求
            print(res["result"])
        else:
            print("%s:%s" % (res["error_code"], res["reason"]))
    else:
        print("request api error")
if __name__ == '__main__':
    main()
```

Step 5: embed your API code to your system

- Question:

Flask can provide us a web interface. Can we embed the code to Flask website?

Results:



```
← → ↻ 127.0.0.1:5000/NewsTouTiao
yule新闻: [{'stat': '1', 'data': [{'uniquekey': '8dcef592397d99a00b4dda5cb9242bd3', 'title': '靳东带来新剧《如果岁月可回头》，女主是《欢乐颂》里的她！', 'date': '2018-11-25 23:39', 'category': '娱乐', 'author_name': '小侠侃娱乐', 'url': 'http://mini.eastday.com/mobile/181125233956738.html', 'thumbnail_pic_s': 'http://00imgmini.eastday.com/mobile/20181125/20181125233956_babef2c05b25753866973d7c12b6509a_3_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://00imgmini.eastday.com/mobile/20181125/20181125233956_babef2c05b25753866973d7c12b6509a_5_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://00imgmini.eastday.com/mobile/20181125/20181125233956_babef2c05b25753866973d7c12b6509a_1_mwpm_03200403.jpg'}, {'uniquekey': 'ab60854edb80d4dbdec84b4109590e5f', 'title': '原来我们所知道的他竟然是蔡明的儿子！网友：果然是隐藏的富二代', 'date': '2018-11-25 23:39', 'category': '娱乐', 'author_name': '娱乐大表姐', 'url': 'http://mini.eastday.com/mobile/1811252339444889.html', 'thumbnail_pic_s': 'http://06imgmini.eastday.com/mobile/20181125/20181125233944_d41d8cd98f00b204e9800998ecf8427e_3_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://06imgmini.eastday.com/mobile/20181125/20181125233944_d41d8cd98f00b204e9800998ecf8427e_4_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://06imgmini.eastday.com/mobile/20181125/20181125233944_d41d8cd98f00b204e9800998ecf8427e_1_mwpm_03200403.jpg'}, {'uniquekey': '4bc8aadcd038aac5a0c464d2ecf8769c8', 'title': '被称为行走的荷尔蒙，实力派小生，老婆被喷丑他霸气回应', 'date': '2018-11-25 23:39', 'category': '娱乐', 'author_name': '趣事推荐', 'url': 'http://mini.eastday.com/mobile/181125233926629.html', 'thumbnail_pic_s': 'http://04imgmini.eastday.com/mobile/20181125/20181125233926_d41d8cd98f00b204e9800998ecf8427e_5_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://04imgmini.eastday.com/mobile/20181125/20181125233926_d41d8cd98f00b204e9800998ecf8427e_1_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://04imgmini.eastday.com/mobile/20181125/20181125233926_d41d8cd98f00b204e9800998ecf8427e_3_mwpm_03200403.jpg'}, {'uniquekey': '808cfcd1d3ae384d72cf77219697dbb', 'title': '《将夜》“三痴”相继上线，只论颜值你喜欢哪个？', 'date': '2018-11-25 23:39', 'category': '娱乐', 'author_name': '娱乐传真机', 'url': 'http://mini.eastday.com/mobile/181125233922512.html', 'thumbnail_pic_s': 'http://07imgmini.eastday.com/mobile/20181125/20181125233922_cb0ca12429b25300596e56bde2978cbc_6_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://07imgmini.eastday.com/mobile/20181125/20181125233922_cb0ca12429b25300596e56bde2978cbc_1_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://07imgmini.eastday.com/mobile/20181125/20181125233922_cb0ca12429b25300596e56bde2978cbc_5_mwpm_03200403.jpg'}, {'uniquekey': '7f72ee2f48067d60108bfc30fc694c4d', 'title': '曝出梁洁不想拍《双世宠妃3》，得知原因后，网友表示：我们理解', 'date': '2018-11-25 23:39', 'category': '娱乐', 'author_name': '娱乐汇', 'url': 'http://mini.eastday.com/mobile/181125233907985.html', 'thumbnail_pic_s': 'http://00imgmini.eastday.com/mobile/20181125/20181125233907_da8c9fdb34120422cf10d1d4166ac05_4_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://00imgmini.eastday.com/mobile/20181125/20181125233907_da8c9fdb34120422cf10d1d4166ac05_1_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://00imgmini.eastday.com/mobile/20181125/20181125233907_da8c9fdb34120422cf10d1d4166ac05_3_mwpm_03200403.jpg'}, {'uniquekey': 'e8ce5eb719312f69f233d4622a74a305', 'title': '杜若溪谈《演员》细节，严屹宽：2天瘦5斤', 'date': '2018-11-25 23:38', 'category': '娱乐', 'author_name': '明星娱乐八卦', 'url': 'http://mini.eastday.com/mobile/181125233852538.html', 'thumbnail_pic_s': 'http://02imgmini.eastday.com/mobile/20181125/20181125233852_d41d8cd98f00b204e9800998ecf8427e_5_mwpm_03200403.jpg', 'thumbnail_pic_s02': 'http://02imgmini.eastday.com/mobile/20181125/20181125233852_d41d8cd98f00b204e9800998ecf8427e_6_mwpm_03200403.jpg', 'thumbnail_pic_s03': 'http://02imgmini.eastday.com/mobile/20181125/20181125233852_d41d8cd98f00b204e9800998ecf8427e_10_mwpm_03200403.jpg'}, {'uniquekey':
```

Web API

```
from flask import Flask
from flask import request
import json, urllib.request
from urllib.parse import urlencode
```

```
app = Flask(__name__)
```

```
@app.route('/', methods=['GET', 'POST'])
```

```
def home():
```

```
    return '''<h1>Home</h1>
    <p><a href="/NewsTouTiao">NewsTouTiao</a></p>'''
```

<-Login

```
@app.route('/NewsTouTiao', methods=['GET'])
```

```
def login_form():
```

```
    return '''<form action="/NewsTouTiao" method="POST">
    <p>Type:
    <select name="NewsType">
    <option value="top">头条</option>
    <option value="shehui">社会</option>
    <option value="guonei">国内</option>
    <option value="guoji">国际</option>
    <option value="yule">娱乐</option>
    <option value="tiyu">体育</option>
    <option value="junshi">军事</option>
    <option value="keji">科技</option>
    <option value="caijing">财经</option>
    <option value="shishang">时尚</option>
    </select>
    </p>
    <p><button type="submit">View</button></p>
    </form>'''
```

```
@app.route('/NewsTouTiao', methods=['POST'])
```

```
def login():
```

```
    # 配置您申请的APPKey
```

```
    appkey = "04f*****17c"
```

<-Main()

```
    NewsType = request.form["NewsType"]
```

```
    # 1. 新闻查询
```

```
    return NewsType+"新闻:" +request1(NewsType, appkey, "GET")
```

```
# 新闻查询
```

```
def request1(NewsType, appkey, m="GET"):
```

```
    url = "http://v.juhe.cn/toutiao/index"
```

```
    params = {
```

```
        "type": NewsType, # top(头条, 默认), shehui(社会), guonei(国内), guoji(国际), yule(娱乐), tiyu(体育), junshi(军事), keji(科技), caijing(财经), shishang(时尚)
```

```
        "key": appkey, # 应用APPKEY(应用详细页查询)
```

```
        "dtype": "", # 返回数据的格式, xml或json, 默认json
```

```
    }
```

```
    params = urlencode(params)
```

```
    if m == "GET":
```

```
        f = urllib.request.urlopen("%s%s" % (url, params))
```

```
    else:
```

```
        f = urllib.request.urlopen(url, params)
```

```
    content = f.read()
```

```
    res = json.loads(content)
```

```
    if res:
```

```
        error_code = res["error_code"]
```

```
        if error_code == 0:
```

```
            # 成功请求
```

```
            return str(res["result"])
```

```
        else:
```

```
            return str(res["error_code"])+" "+ str(res["reason"])
```

```
    else:
```

```
        return "request api error"
```

```
if __name__ == '__main__':
```

```
    app.run()
```

<-almost the same



Tips:

1. Register, Login, and get your **API KEY** from each website before you use API.
2. Different websites have different API Keys.
3. Copy and Paste, **Standing on the Shoulders of Giants**





Ask A Question



*What are the **advantages** and **disadvantages** of using API?*

• *Advantages:*

- much easier to get useful information*
- more efficient to develop website platforms*
- more helps from professionals*

• *Disadvantages:*

- maybe not free*
- maybe limitations*
(cannot use it sometime in the future)
- maybe some mistakes or exceptions*



human-readable text for data transmission

JSON

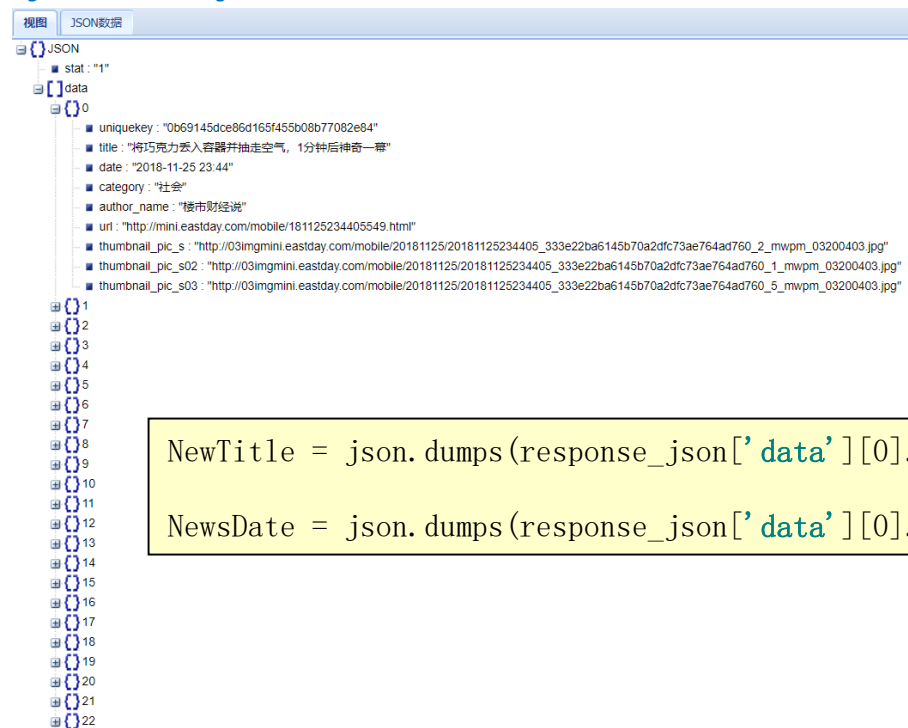
Analyzing JSON

1. `import json`
2. `json.loads`
3. `json.dumps`
4. get data following the tree
5. output results



JSON结构视图

<http://www.bejson.com/jsonviewernew/>



有中文

```
NewTitle = json.dumps(response_json['data'][0].get("title"), ensure_ascii=False)

NewsDate = json.dumps(response_json['data'][0].get("date"))
```

无中文

- The result :

```
{‘stat’: ‘1’, ‘data’: [{‘uniquekey’: ‘0b69145dce86d165f455b08b77082e84’,  
‘title’: ‘将巧克力丢入容器并抽走空气，1分钟后神奇一幕’, ‘date’:  
‘2018-11-25 23:44’, ‘category’: ‘社会’, ‘author_name’: ‘楼市财经说’,  
‘url’: ‘http://mini.eastday.com/mobile/181125234405549.html’,  
‘thumbnail_pic_s’:  
‘http://03imgmini.eastday.com/mobile/20181125/20181125234405_333e22  
ba6145b70a2dfc73ae764ad760_2_mwpm_03200403.jpg’,  
‘thumbnail_pic_s02’:  
‘http://03imgmini.eastday.com/mobile/20181125/20181125234405_333e22  
ba6145b70a2dfc73ae764ad760_1_mwpm_03200403.jpg’,  
‘thumbnail_pic_s03’:  
‘http://03imgmini.eastday.com/mobile/20181125/20181125234405_333e22  
ba6145b70a2dfc73ae764ad760_5_mwpm_03200403.jpg’}],.....}
```

JSON

```
@app.route('/NewsTouTiao', methods=['POST'])
def login():
    # 配置您申请的APPKey
    appkey = "04f*****17c"
    NewsType = request.form["NewsType"]
    if NewsType == "top":
        CNewsType = "头条"
    elif NewsType == "shehui":
        CNewsType = "社会"
    elif NewsType == "guonei":
        CNewsType = "国内"
    elif NewsType == "guoji":
        CNewsType = "国际"
    elif NewsType == "yule":
        CNewsType = "娱乐"
    elif NewsType == "tiyu":
        CNewsType = "体育"
    elif NewsType == "junshi":
        CNewsType = "军事"
    elif NewsType == "keji":
        CNewsType = "科技"
    elif NewsType == "caijing":
        CNewsType = "财经"
    elif NewsType == "shishang":
        CNewsType = "时尚"
    else:
        CNewsType = "头条"
    ResultTable = "<table><tr><td>#</td><td>Title</td><td>Date</td><td>Category</td><td>Author</td></tr>"
    response = request1(NewsType, appkey, "GET").replace("'", "")
    response_json = json.loads(response)
    for i in range(30):
        NewTitle = json.dumps(response_json['data'][i].get("title"), ensure_ascii=False)
        NewsDate = json.dumps(response_json['data'][i].get("date"))
        NewCate = json.dumps(response_json['data'][i].get("category"), ensure_ascii=False)
        NewAuthor = json.dumps(response_json['data'][i].get("author_name"), ensure_ascii=False)
        NewsURL = json.dumps(response_json['data'][i].get("url"))
        ResultTable += "<tr><td>"+str((i+1))+"</td><td><a href='"+NewsURL[1:-1]+" target='_blank'>"+NewTitle[1:-1]+
1)+"</a></td><td>"+NewsDate[1:-1]+"</td><td>"+NewCate[1:-1]+"</td><td>"+NewAuthor[1:-1]+"</td></tr>"
    ResultTable += "</table>"
    # 1. 新闻查询
    return CNewsType+"新闻:" + ResultTable
```

EXAMPLE 2: News 2

NEWS

Example

← → ↺ 127.0.0.1:5000/NewsTouTiao

财经新闻:

Title

- 1 [实体店定价超原产国 加拿大鹅打击代购缺诚意](#)
- 2 [当U盘已成往事：关于青春，你还记得什么？](#)
- 3 [家庭不和睦，人就会生病！](#)
- 4 [“爸，在忙吗？”这两天，杭州多位家长收到这种短信](#)
- 5 [捐款1个亿！80后夫妇王麒诚和妻子吴艳浙大演讲，满满的创业干货](#)
- 6 [企业出现如下几种现象，仍不进行变革，结局只能是关门大吉](#)
- 7 [东北玉米价位已至高位 或难以持续上涨行情](#)
- 8 [大蒜市场观察：今年大蒜市场遇冷，2019年会上涨吗？](#)
- 9 [这个一线城市二手房市场凉凉，业主降价15%卖房，中介半年没接单](#)
- 10 [龙头点评：目前市场唯一四板，有望领涨大盘（附股）](#)
- 11 [再生水利用将成破解水资源短缺重要途径 千亿级水处理行业迎来新风口](#)
- 12 [四川自贸试验区：为内陆开放添动力](#)
- 13 [国际煤市要闻回顾（11.19-11.23）](#)
- 14 [目前主产区玉米价格仍处于上涨通道中 农民怎么才能卖个最高价](#)
- 15 [家家悦亿元入股福悦祥，11家门店及13万平物流中心尽归麾下！](#)
- 16 [柬埔寨智慧农业领域发展潜力依旧巨大](#)
- 17 [300多天，400余次调控，未来房价如何走](#)
- 18 [实话告诉你！未来10年才是最好的买房时期](#)
- 19 [房价涨愁，房价跌忧，刚需到底怎么了？](#)
- 20 [这6只年报预增翻倍股被机构盯上了！](#)
- 21 [马云讲：没钱该不该创业三句话选择创业新思路](#)
- 22 [股价“梦回2017”，缺钱的复星医药正“骑虎难下”](#)
- 23 [家族信托与集合资金信托的区别是什么？](#)
- 24 [从禁止石油开采到成为全球头号产油国，美国为何突然大量出口石油](#)

Date

Category Author

- | | | |
|------------------|----|--------------|
| 2018-11-26 00:29 | 财经 | 北京商报 |
| 2018-11-26 00:27 | 财经 | 中国新闻网 |
| 2018-11-26 00:07 | 财经 | 一生致富 |
| 2018-11-25 23:49 | 财经 | 杭州卡洛儿 |
| 2018-11-25 23:46 | 财经 | GW34568创业家 |
| 2018-11-25 23:24 | 财经 | 诺言说职场 |
| 2018-11-25 23:23 | 财经 | 科学兴农 |
| 2018-11-25 23:23 | 财经 | 科学兴农 |
| 2018-11-25 23:19 | 财经 | 每日经济新闻 |
| 2018-11-25 23:19 | 财经 | 美湖路的老程 |
| 2018-11-25 23:17 | 财经 | 每日经济新闻 |
| 2018-11-25 23:17 | 财经 | 新华网 |
| 2018-11-25 22:48 | 财经 | 懂焦帝 |
| 2018-11-25 22:45 | 财经 | 每日粮油 |
| 2018-11-25 22:41 | 财经 | 第一纺织网 |
| 2018-11-25 22:40 | 财经 | 华彩国际传媒 |
| 2018-11-25 22:27 | 财经 | 今朝聊房事 |
| 2018-11-25 22:27 | 财经 | 深度金融 |
| 2018-11-25 22:27 | 财经 | 天狼说地产 |
| 2018-11-25 22:25 | 财经 | 21世纪经济报道 |
| 2018-11-25 22:24 | 财经 | 要树立干事创业的鲜明导向 |
| 2018-11-25 22:23 | 财经 | 中访网 |
| 2018-11-25 22:17 | 财经 | 超仁博一博 |
| 2018-11-25 22:16 | 财经 | 环球财经天下 |





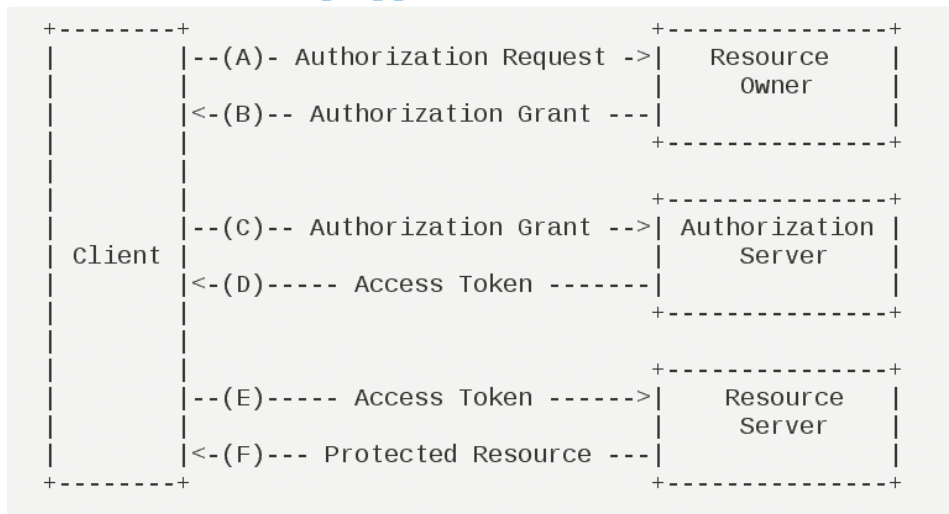
some examples on social media

Web API Applications

Web API Applications

OAuth

An open protocol to allow secure API authorization in a simple and standard method from web, mobile and desktop applications.



API Key
needed

Access Token
needed



Ref. http://www.ruanyifeng.com/blog/2014/05/oauth_2_0.html

Web API Applications

EXAMPLE 3: Friendship of TVs on Twitter



Friendship of TVs on Twitter

Problem Description

This research aims to investigate the significance and activeness of TV media companies in the world. Twitter friendships between TV broadcasting companies were employed for this research. Totally, about 104 TV channels from different countries are selected for this research. In-Degree and Out-Degree are used as features in this study.

Features:

In-Degree: Followed by Others, which shows the significance to the world

Out-Degree: Following Others, which shows the activeness in the world



Friendship of TVs on Twitter

twitter and Its API

twitter

Do NOT tell me you do NOT know twitter, even you are a Chinese.

If you really do not know it, visit <https://twitter.com> right now!



twitter API: <https://dev.twitter.com/>



Friendship of TVs on Twitter

Tweepy

An easy-to-use Python library for accessing the Twitter API

Installation

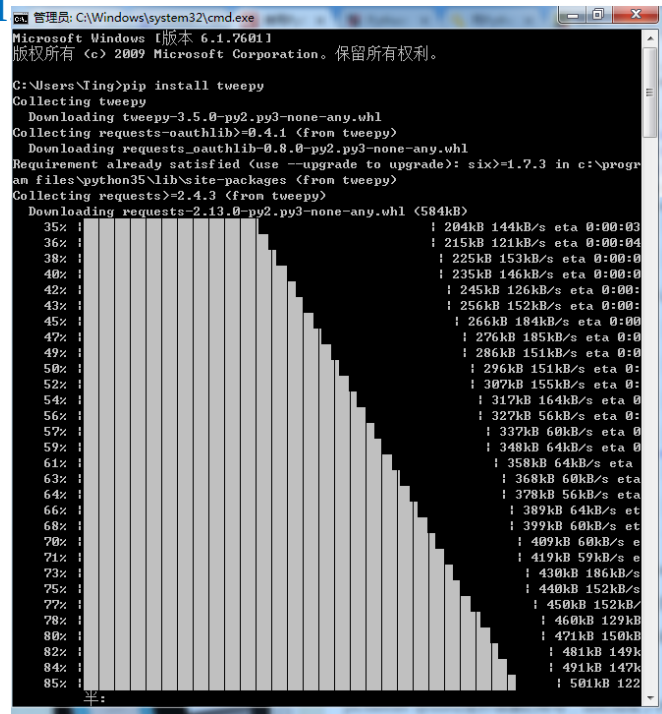
pip install tweepy

Reference to Tweepy

<https://github.com/tweepy/tweepy>

<http://www.tweepy.org/>

<http://pythonhosted.org/tweepy/>



```
管理员: C:\Windows\system32\cmd.exe
Microsoft Windows [版本 6.1.7601]
版权所有 (c) 2009 Microsoft Corporation. 保留所有权利。

C:\Users\Ting>pip install tweepy
Collecting tweepy
  Downloading tweepy-3.5.0-py2.py3-none-any.whl
Collecting requests-oauthlib-0.4.1 (from tweepy)
  Downloading requests_oauthlib-0.8.0-py2.py3-none-any.whl
Requirement already satisfied (use --upgrade to upgrade): six>=1.7.3 in c:\program files\python35\lib\site-packages (from tweepy)
Collecting requests>=2.4.3 (from tweepy)
  Downloading requests-2.13.0-py2.py3-none-any.whl (584kB)
35% |#####| 204kB 144kB/s eta 0:00:03
36% |#####| 215kB 121kB/s eta 0:00:04
38% |#####| 225kB 153kB/s eta 0:00:00
40% |#####| 235kB 146kB/s eta 0:00:00
42% |#####| 245kB 126kB/s eta 0:00:00
43% |#####| 256kB 152kB/s eta 0:00:00
45% |#####| 266kB 184kB/s eta 0:00:00
47% |#####| 276kB 185kB/s eta 0:00:00
49% |#####| 286kB 151kB/s eta 0:00:00
50% |#####| 296kB 151kB/s eta 0:00:00
52% |#####| 307kB 155kB/s eta 0:00:00
54% |#####| 317kB 164kB/s eta 0:00:00
56% |#####| 327kB 56kB/s eta 0:00:00
57% |#####| 337kB 60kB/s eta 0:00:00
59% |#####| 348kB 64kB/s eta 0:00:00
61% |#####| 358kB 64kB/s eta 0:00:00
63% |#####| 368kB 60kB/s eta 0:00:00
64% |#####| 378kB 56kB/s eta 0:00:00
66% |#####| 389kB 64kB/s eta 0:00:00
68% |#####| 399kB 60kB/s eta 0:00:00
70% |#####| 409kB 60kB/s eta 0:00:00
71% |#####| 419kB 59kB/s eta 0:00:00
73% |#####| 430kB 186kB/s eta 0:00:00
75% |#####| 440kB 152kB/s eta 0:00:00
77% |#####| 450kB 152kB/s eta 0:00:00
78% |#####| 460kB 129kB/s eta 0:00:00
80% |#####| 471kB 150kB/s eta 0:00:00
82% |#####| 481kB 149kB/s eta 0:00:00
84% |#####| 491kB 147kB/s eta 0:00:00
85% |#####| 501kB 122kB/s eta 0:00:00
```



Friendship of TVs on Twitter

Register Your Twitter APPs

Step 1 : visit <https://apps.twitter.com/>, and click on the button “Create New APP”.

 Application Management

Twitter Apps

You don't currently have any Twitter Apps.

Create New App



Friendship of TVs on Twitter

Step2 : Input the detailed information of your new app.

Create an application

Application Details

Name *

Your application name. This is used to attribute the source of a tweet and in user-facing authorization screens. 32 characters max.

Description *

Your application description, which will be shown in user-facing authorization screens. Between 10 and 200 characters max.

Website *

Your application's publicly accessible home page, where users can go to download, make use of, or find out more information about your application. This fully-qualified URL is used in the source attribution for tweets created by your application and will be shown in user-facing authorization screens.

(If you don't have a URL yet, just put a placeholder here but remember to change it later.)

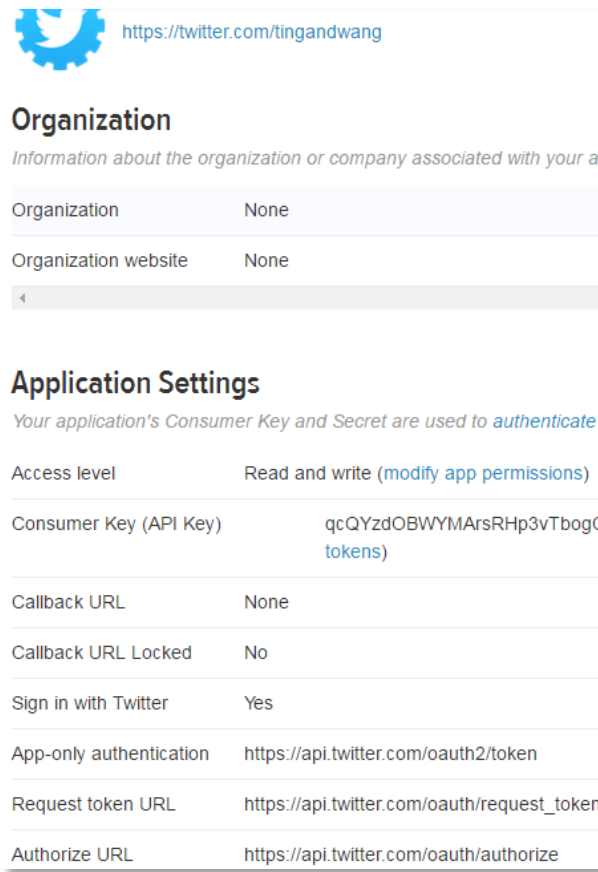
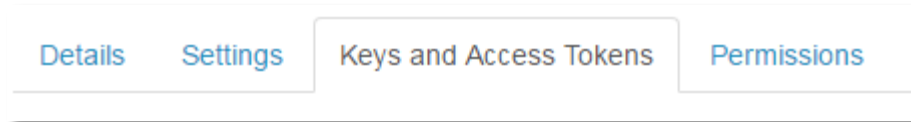


Friendship of TVs on Twitter

Step3: Get your API keys and Access Tokens

[Notes]

You may get all these information when you click on “**Keys and Access Tokens**”.



Friendship of TVs on Twitter

Steps for the System

1. Connect to twitter by API Keys and Access Tokens
2. Find the relationship using Tweepy functions

```
API.show_friendship(source_id/source_screen_name, target_id/target_screen_name)
```

Returns detailed information about the relationship between two users.

- Parameters:
- **source_id** – The user_id of the subject user.
 - **source_screen_name** – The screen_name of the subject user.
 - **target_id** – The user_id of the target user.
 - **target_screen_name** – The screen_name of the target user.

Return type: `Friendship` object

3. Save the results



Friendship of TVs on Twitter

Results

In-Degree for the significance of TV media

In-Degree Ranking

	Country	TV Media	Twitter ID	In-Degree
1	USA	CNN	CNN	26
2	Russia	Russia Today	RT_com	14
3	UK	BBC	BBCNews	13
4	Qatar	Al Jazeera television	AJENews	6
5	Spain	Radio Televisión Española	rtve	5
6	Netherland	Nederlandse Omroep Stichting	NOS	2
7	South Africa	South African Broadcasting Corporation, SABC	SABCNewsOnline	2
8	Lebanon	Al-Manar TV	almanarnews	2



Friendship of TVs on Twitter

Out-Degree for the activeness of TV media

Out-Degree Ranking

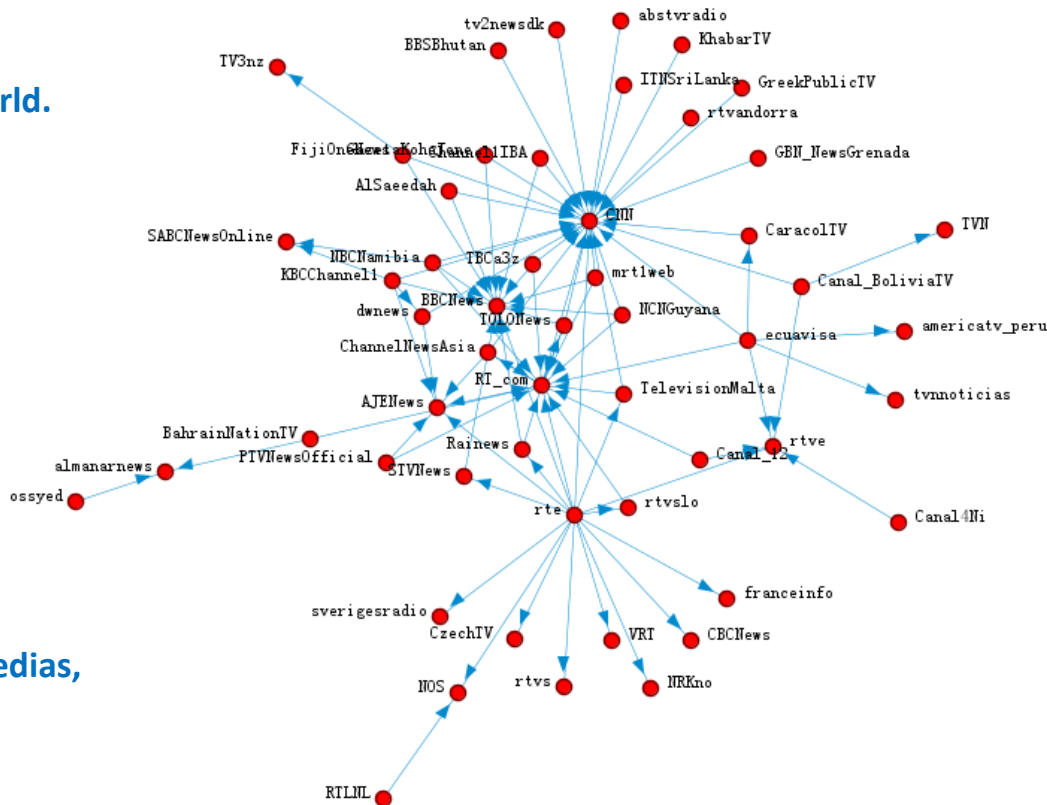
	Country	TV Media	Twitter ID	Out-Degree
1	Ireland	Raidió Teilifís Éireann, Ireland's National Public Service Multi-Media Organisation	rte	17
2	Ecuador	Ecuavisa	ecuavisa	6
3	Kenya	Kenya Broadcasting Corporation	KBCChannel1	5
4	Namibia	Namibian Broadcasting Corporation	NBCNamibia	4
5	Russia	Russia Today	RT_com	4
6	Bolivia	Bolivia TV	Canal_BoliviaTV	3
7	Singapore	Television Corporation of Singapore	ChannelNewsAsia	3
8	Fiji	Fiji Television Limited	FijiOneNews	3
9	Guyana	National Communications Network, Guyana	NCNGuyana	3
10	Tonga	Tonga Broadcasting Commission	TBCa3z	3
11	Afghanistan	TOLONews	TOLONews	3
12	Macedonia	Macedonian Radio-Television	mrt1web	3



Friendship of TVs on Twitter

Conclusions

1. CNN, BBCNews, and Russia Today have the greatest significance to the world.
2. Raidió Teilifís Éireann from Ireland is the most active TV media.
3. Media from developed countries have greater influence.
4. Media from developing countries are more active.
5. CCTV, China neither follow any other medias, nor be followed by any other medias.

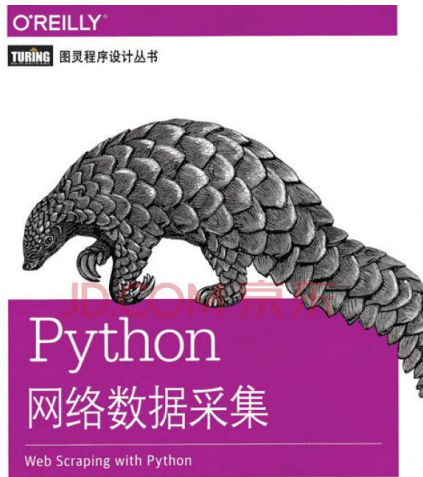




上海外国语大学
SHANGHAI INTERNATIONAL STUDIES UNIVERSITY

Reference

- XML
 - <http://www.w3school.com.cn/x.asp>
- JSON
 - <http://www.w3school.com.cn/json>



[美] Ryan Mitchell 著
陶俊杰 陈小莉 译

中国工信出版集团 人民邮电出版社
CHINA TELECOM PRESS PUBLISHING GROUP PUBLISHING & TELECOM PRESS



Reference

Baidu	(http://apistore.baidu.com/)
Weibo	(http://open.weibo.com/wiki/微博API)
WeChat	(http://dev.wechat.com/)
QQ Tencent	(http://wiki.open.qq.com/wiki/API文档)
360	(http://open.360.cn/)
Taobao	(http://open.taobao.com/docs/api_list.htm?spm=a219a.7386653.1.22.3XIaLs)
JD	(http://jos.jd.com/)
Suning	(http://open.suning.com/)
Douban	(https://developers.douban.com/)
Juhe	(https://www.juhe.cn/)
Google	(https://developers.google.com/)
Facebook	(http://developers.facebook.com/)
Twitter	(https://dev.twitter.com/)
Linkedin	(https://developer.linkedin.com/)





上海外国语大学
SHANGHAI INTERNATIONAL STUDIES UNIVERSITY

Homework

Homework

1. Save the results of Example 2 to Data Base.
Deadline: April. 29th.

Please DO NOT forget to send your code to
tingwang_sisu@163.com





The End of Lecture 7

Thank You

<http://www.wangting.ac.cn>

