

实现支持微信小程序的 聊天机器人

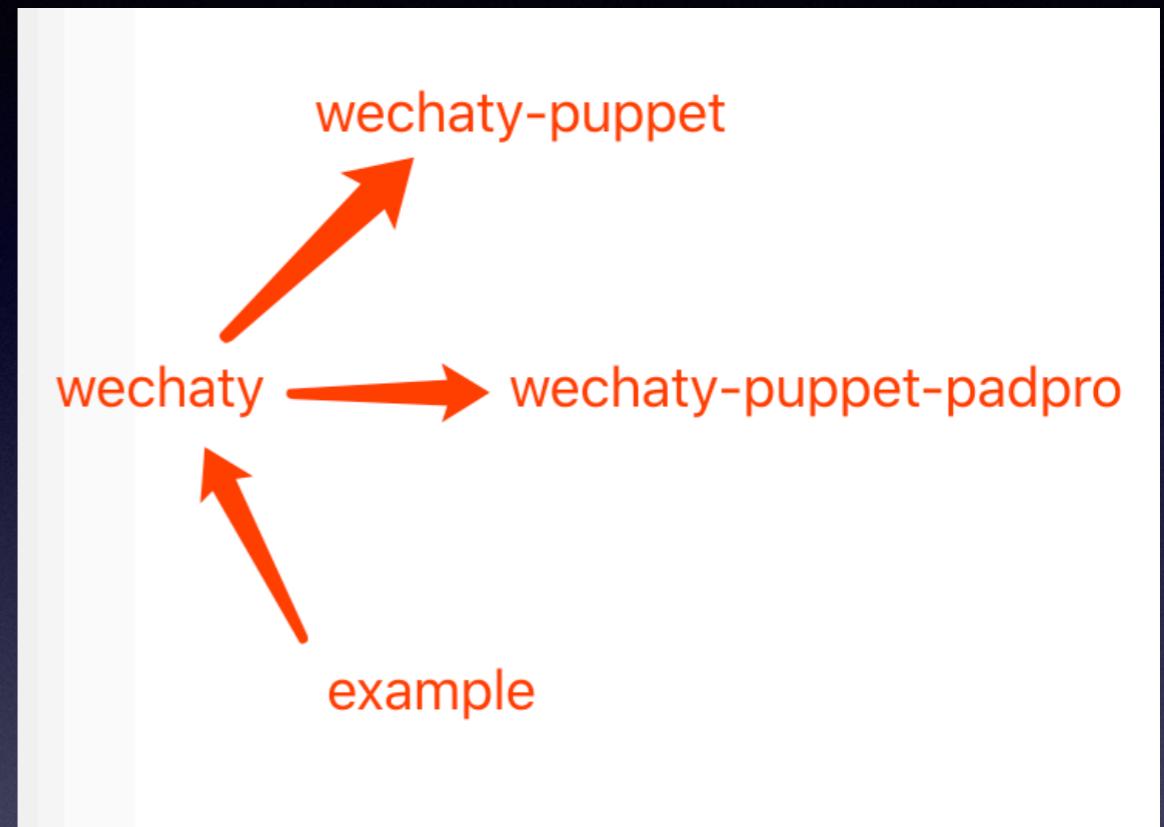
赵俊良
2019.8.16

简介

- wechaty小程序的实现
- 用正则表达式实现聊天机器人
- demo

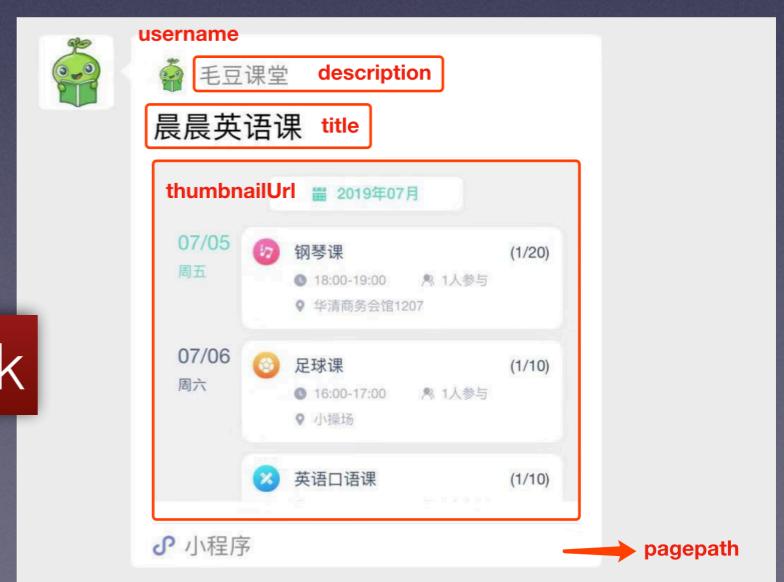
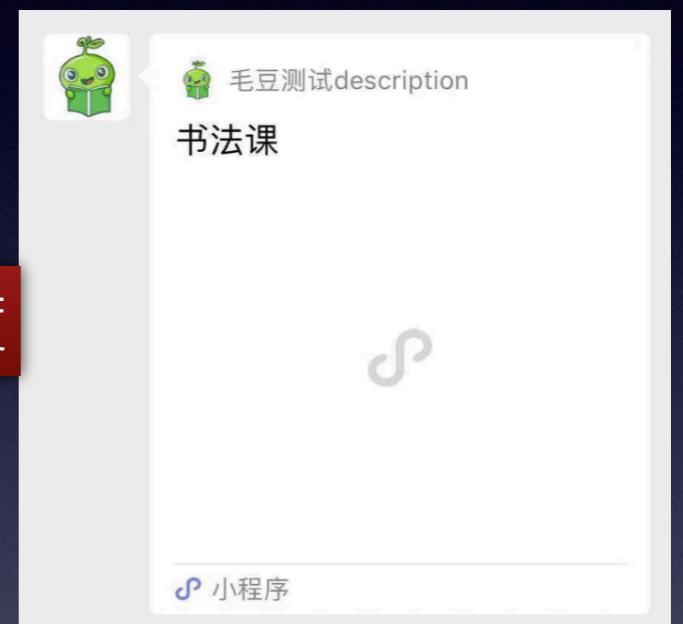
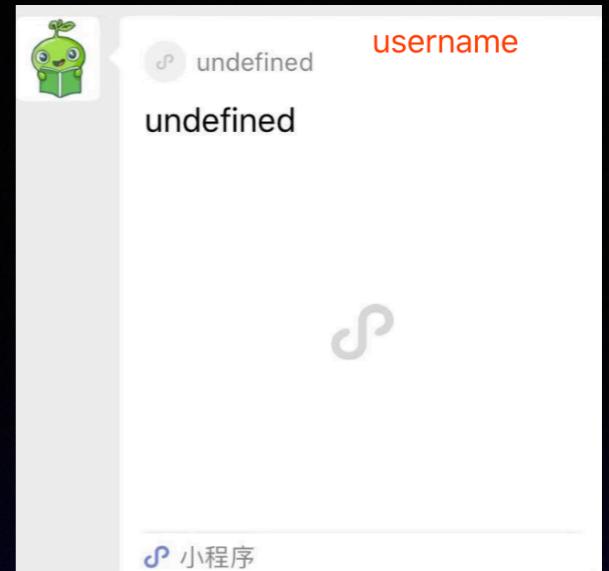
wechaty的开发

- 本地运行开发环境的建立
- 三个库的依赖关系



测试

- 只有username, 可以发出 目前进度
- 图片还不能生成
- 可通过解析小程序, 找到相关信息 hack



wechaty

- 接口层，各种say
- contact.say()
- room.say()

wechaty-puppet

- 抽象层，定义

```
export interface MiniProgramPayload {  
    appid? : string,    // optional, appid, get from wechat (mp.weixin.qq.com)  
    description? : string,    // optional, mini program title  
    pagepath? : string,    // optional, mini program page path  
    thumbnailurl? : string,    // optional, default picture, convert to thumbnail  
    title? : string,    // optional, mini program title  
    username? : string,    // original ID, get from wechat (mp.weixin.qq.com)  
}
```

```
public abstract async messageMiniProgram (messageId: string) : Promise<MiniProgramPayload>  
public abstract async messageSendMiniProgram (receiver: Receiver, miniProgramPayload: MiniProgramPayload) : Promise<void>
```

wechaty-puppet-padpro

- 实现层
- 具体的实现

自然语言识别

- 主要目的：识别时间
- 使用微软@microsoft/recognizers-text-suite
- 识别顺序

```
datetimeV2.datetime  
datetimeV2.time  
datetimeV2.datetimerange  
datetimeV2.date  
| datetimeV2.daterange  
| datetimeV2.datetimerange
```

去除特殊符号

- 去表情

```
var msgText = originalText.replace(/<[^>]*>?/gm, '')
```

- 去开头空格

```
msgText = msgText.replace(/^\s*/g, '')
```

时间歧义1

The screenshot shows a WeChat message from a bot. The message content is: "周日晚上6:30, 去bot5沙龙". A red box highlights this text. Below it is a green button with the same text. To the right is a small sun icon.

[课程提醒创建成功通知]

标题: 去沙龙
时间: 8/18/2019, 6:30:00 PM
地点: bot5沙龙

消息原文:
周日晚上6:30, 去bot5沙龙

[你收到了一个小程序, 请在手机上查看]

At the bottom, there are standard WeChat interaction icons: smiley face, file, share, and more.

On the right side of the screen, a terminal window displays the JSON log of the NLP processing:

```
text: '周日晚上6:30',
typeName: 'datetimeV2.datetime' } ] +25ms
aodou-nlp.js {
aodou-nlp.js      "start": 0,
aodou-nlp.js      "end": 7,
aodou-nlp.js      "resolution": {
aodou-nlp.js        "values": [
aodou-nlp.js          {
aodou-nlp.js            "timex": "XXXX-WXX-7T18:30",
aodou-nlp.js            "type": "datetime",
aodou-nlp.js            "value": "2019-08-11 18:30:00"
aodou-nlp.js        },
aodou-nlp.js        {
aodou-nlp.js          "timex": "XXXX-WXX-7T18:30",
aodou-nlp.js          "type": "datetime",
aodou-nlp.js          "value": "2019-08-18 18:30:00"
aodou-nlp.js      ],
aodou-nlp.js      "text": "周日晚上6:30",
aodou-nlp.js      "typeName": "datetimeV2.datetime"
aodou-nlp.js } +1ms
```

时间歧义2

The screenshot shows a WeChat message from a bot. The message content is "6月9日10点, 我也参加了wechaty沙龙". A red box highlights this timestamp. Below it is a card with course reminder information: Title: 参加沙龙, Time: 6/9/2019, 10:00:00 PM, Location: 沙龙. Another card says "你收到了一个小程序, 请在手机上查看". At the bottom are standard WeChat message controls.

```
aodou-nlp.js      "end": 6,
aodou-nlp.js      "resolution": {
aodou-nlp.js        "values": [
aodou-nlp.js          {
aodou-nlp.js            "timex": "XXXX-06-09T10",
aodou-nlp.js            "type": "datetime",
aodou-nlp.js            "value": "2019-06-09 10:00:00"
aodou-nlp.js        },
aodou-nlp.js      },
aodou-nlp.js      "timex": "XXXX-06-09T22",
aodou-nlp.js      "type": "datetime",
aodou-nlp.js      "value": "2019-06-09 22:00:00"
aodou-nlp.js    ],
aodou-nlp.js    },
aodou-nlp.js    "timex": "XXXX-06-09T10",
aodou-nlp.js    "type": "datetime",
aodou-nlp.js    "value": "2020-06-09 10:00:00"
aodou-nlp.js  },
aodou-nlp.js  },
aodou-nlp.js  "timex": "XXXX-06-09T22",
aodou-nlp.js  "type": "datetime",
aodou-nlp.js  "value": "2020-06-09 22:00:00"
aodou-nlp.js ]
aodou-nlp.js  },
aodou-nlp.js    "text": "6月9日10点",
aodou-nlp.js    "typeName": "datetimeV2.datetime"
aodou-nlp.js } +0ms
aodou-nlp.js getTimeInResult() { results:
aodou-nlp.js  f.starts[0]
```

时间歧义3

The screenshot shows a WeChat conversation. A message from a contact (represented by a green icon) contains the text: "6月9日22点, 我也参加了wechaty沙龙". This message is highlighted with a red rectangle. Below it, a message from the user says: "[你收到了一个小程序, 请在手机上查看]". Another message from the contact says: "[课程提醒创建成功通知]". Inside this message, there is a detailed description of an event: "标题: 参加沙龙", "时间: 6/9/2020, 10:00:00 PM", and "地点: 沙龙". At the bottom of the conversation, there are standard WeChat interaction icons: smiley face, file, crossed-out file, and three dots.

```
resolution: { values: [Array] },
text: '6月9日22点',
typeName: 'datetimeV2.datetime' } ] +217ms
maodou-nlp.js {
maodou-nlp.js   "start": 0,
maodou-nlp.js   "end": 6,
maodou-nlp.js   "resolution": {
maodou-nlp.js     "values": [
maodou-nlp.js       {
maodou-nlp.js         "timex": "XXXX-06-09T22",
maodou-nlp.js         "type": "datetime",
maodou-nlp.js         "value": "2019-06-09 22:00:00"
},
{
maodou-nlp.js         "timex": "XXXX-06-09T22",
maodou-nlp.js         "type": "datetime",
maodou-nlp.js         "value": "2020-06-09 22:00:00"
}
],
maodou-nlp.js   },
maodou-nlp.js   "text": "6月9日22点",
maodou-nlp.js   "typeName": "datetimeV2.datetime"
maodou-nlp.js } +0ms
```

```
var msgText2 = msgText.replace(/号楼/g, '#号楼')
```

九号楼歧义

【暑期特别策划】郑渊洁长篇故事——《魔方大厦》
活动地点：幸福城儿童俱乐部（幸福城欣园九号楼二单元101）
活动时间：7月14日周日上午 10:00-11:30
报名接龙：群内所有小朋友（4-12岁），会员优先
报名限额：12名小朋友
123Lab幸福城儿童俱乐部故事会

[课程提醒创建成功通知]

标题：策划故事魔方
时间：9/9/2019, 10:00:00 AM
地点：幸福城欣园九号楼二单元101

消息原文：

@所有人
【暑期特别策划】郑渊洁长篇故事——《魔方大厦》
活动地点：幸福城儿童俱乐部（幸福城欣园九号楼二单元101）
活动时间：7月14日周日上午 10:00-11:30
报名接龙：群内所有小朋友（4-12岁），会员优先
报名限额：12名小朋友
123Lab幸福城儿童俱乐部故事会

```
resolution: { values: [Array] },  
text: '10:00-11:30',  
typeName: 'datetimeV2.timerange' } ] +74ms  
maodou-nlp.js {  
maodou-nlp.js "start": 48,  
maodou-nlp.js "end": 49,  
maodou-nlp.js "resolution": {  
maodou-nlp.js "values": [  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-XX-09",  
maodou-nlp.js "type": "date",  
maodou-nlp.js "value": "2019-08-09"  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-XX-09",  
maodou-nlp.js "type": "date",  
maodou-nlp.js "value": "2019-09-09"  
maodou-nlp.js ]  
maodou-nlp.js },  
maodou-nlp.js "text": "九号",  
maodou-nlp.js "typeName": "datetimeV2.date"  
maodou-nlp.js } +1ms  
maodou-nlp.js {  
maodou-nlp.js "start": 64,  
maodou-nlp.js "end": 72.  
maodou-nlp.js {  
maodou-nlp.js "start": 65,  
maodou-nlp.js "end": 73,  
maodou-nlp.js "resolution": {  
maodou-nlp.js "values": [  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-07-14TMO",  
maodou-nlp.js "type": "datetimerange",  
maodou-nlp.js "start": "2019-07-14 08:00:00",  
maodou-nlp.js "end": "2019-07-14 12:00:00"  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-07-14TMO",  
maodou-nlp.js "type": "datetimerange",  
maodou-nlp.js "start": "2020-07-14 08:00:00"  
maodou-nlp.js "end": "2020-07-14 12:00:00"  
maodou-nlp.js ]  
maodou-nlp.js },  
maodou-nlp.js "text": "7月14日周日上午",  
maodou-nlp.js "typeName": "datetimeV2.datetimerange"  
maodou-nlp.js }
```

[课程提醒创建成功通知]

标题：策划故事魔方
时间：7/14/2020, 10:00:00 AM
地点：幸福城欣园九号楼二单元101

消息原文：

@所有人
【暑期特别策划】郑渊洁长篇故事——《魔方大厦》
活动地点：幸福城儿童俱乐部（幸福城欣园九号楼二单元101）

```
resolution: { values: [Array] },  
text: '10:00-11:30',  
typeName: 'datetimeV2.timerange' } ] +74ms  
maodou-nlp.js {  
maodou-nlp.js "start": 48,  
maodou-nlp.js "end": 49,  
maodou-nlp.js "resolution": {  
maodou-nlp.js "values": [  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-XX-09",  
maodou-nlp.js "type": "date",  
maodou-nlp.js "value": "2019-08-09"  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-XX-09",  
maodou-nlp.js "type": "date",  
maodou-nlp.js "value": "2019-09-09"  
maodou-nlp.js ]  
maodou-nlp.js },  
maodou-nlp.js "text": "九号",  
maodou-nlp.js "typeName": "datetimeV2.date"  
maodou-nlp.js } +1ms  
maodou-nlp.js {  
maodou-nlp.js "start": 64,  
maodou-nlp.js "end": 72.  
maodou-nlp.js {  
maodou-nlp.js "start": 65,  
maodou-nlp.js "end": 73,  
maodou-nlp.js "resolution": {  
maodou-nlp.js "values": [  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-07-14TMO",  
maodou-nlp.js "type": "datetimerange",  
maodou-nlp.js "start": "2019-07-14 08:00:00",  
maodou-nlp.js "end": "2019-07-14 12:00:00"  
maodou-nlp.js {  
maodou-nlp.js "timex": "XXXX-07-14TMO",  
maodou-nlp.js "type": "datetimerange",  
maodou-nlp.js "start": "2020-07-14 08:00:00"  
maodou-nlp.js "end": "2020-07-14 12:00:00"  
maodou-nlp.js ]  
maodou-nlp.js },  
maodou-nlp.js "text": "7月14日周日上午",  
maodou-nlp.js "typeName": "datetimeV2.datetimerange"  
maodou-nlp.js }
```

```
var msgText2 = msgText2.replace(/(\d+)\-(\d+)点/g, '$1点-$2点')
```

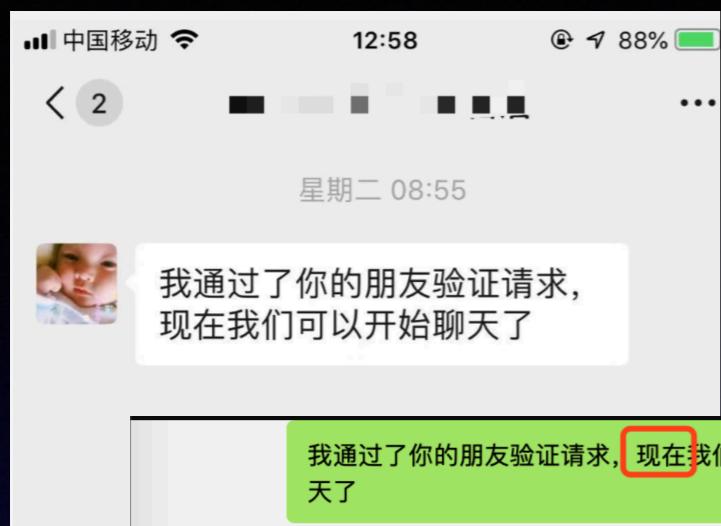
2-6点歧义

The screenshot shows two messages in a messaging application. Both messages contain the same text: "标题: 书法课", "时间: 7/14/2020, 6:00:00 AM", and "地点: 2书法课". The first message was sent at 16:50, and the second message was sent at 16:51. The log output from the `maodou-nlp.js` file shows the processing of these messages, specifically focusing on the date and time resolution for the second message.

Log Output (maodou-nlp.js):

```
maodou-nlp.js      "start": 0,
maodou-nlp.js      "end": 8,
maodou-nlp.js      "resolution": {
maodou-nlp.js          "values": [
maodou-nlp.js      {
maodou-nlp.js          "timex": "XXXX-07-14TAF",
maodou-nlp.js          "type": "datetimerange",
maodou-nlp.js          "start": "2019-07-14 12:00:00",
maodou-nlp.js          "end": "2019-07-14 16:00:00"
},
{
maodou-nlp.js          "timex": "XXXX-07-14TAF",
maodou-nlp.js          "type": "datetimerange",
maodou-nlp.js          "start": "2020-07-14 12:00:00",
maodou-nlp.js          "end": "2020-07-14 16:00:00"
}
],
"text": "7月14日周日下午",
typeNameName": "datetimeV2.datetimerange"
text: '7月14日周日下午2点-6点',
typeName: 'datetimeV2.datetimerange' } ] +40ms
maodou-nlp.js {
maodou-nlp.js      "start": 0,
maodou-nlp.js      "end": 13,
maodou-nlp.js      "resolution": {
maodou-nlp.js          "values": [
maodou-nlp.js      {
maodou-nlp.js          "timex": "(XXXX-07-14T14,XXXX-07-14T18,PT4H)",
maodou-nlp.js          "type": "datetimerange",
maodou-nlp.js          "start": "2019-07-14 14:00:00",
maodou-nlp.js          "end": "2019-07-14 18:00:00"
},
{
maodou-nlp.js          "timex": "(XXXX-07-14T14,XXXX-07-14T18,PT4H)",
maodou-nlp.js          "type": "datetimerange",
maodou-nlp.js          "start": "2020-07-14 14:00:00",
maodou-nlp.js          "end": "2020-07-14 18:00:00"
}
],
"text": "7月14日周日下午2点-6点",
typeName": "datetimeV2.datetimerange"
maodou-nlp.js } +1ms
maodou-nlp.js getTimeInResult() { results:
```

```
var msgText2 = msgText2.replace(/现在/g, '')
```



忽略现在

A screenshot of a mobile messaging application. On the left, a message from a friend is shown with a timestamp: "时间: 8/15/2019, 4:58:19 PM". This timestamp is highlighted with a red rectangle. On the right, a terminal window displays the processing of this message by the NLP engine. The log shows the input message, the results after date/time recognition, and the specific datetime object extracted. The datetime object is highlighted with a red rectangle and contains the following details:

```
maodou-nlp.js input before recognizeDateTime 我通过了你的朋友验证请求, 现在我们可以开始聊天了  
maodou-nlp.js results after recognizeDateTime [ { start: 13,  
  end: 14,  
  resolution: { values: [Array] },  
  text: '现在',  
  typeName: 'datetimeV2.datetime' } ] +36ms  
maodou-nlp.js {  
  maodou-nlp.js   "start": 13,  
  maodou-nlp.js   "end": 14,  
  maodou-nlp.js   "resolution": {  
    maodou-nlp.js     "values": [  
      maodou-nlp.js       {  
        maodou-nlp.js         "timex": "PRESENT_REF",  
        maodou-nlp.js         "type": "datetime",  
        maodou-nlp.js         "value": "2019-08-15 16:58:19"  
      }  
    ],  
    maodou-nlp.js   "text": "现在",  
    maodou-nlp.js   "typeName": "datetimeV2.datetime"  
  }  
  maodou-nlp.js } +0ms  
maodou-nlp.js getTimeInResult() { results:
```

A screenshot of a mobile messaging application. On the left, a message from a friend is shown with a timestamp: "时间: 8/15/2019, 4:58:19 PM". This timestamp is highlighted with a red rectangle. On the right, a terminal window displays the processing of this message by the NLP engine. The log shows the input message, the results after date/time recognition, and the specific datetime object extracted. A red arrow points from the highlighted timestamp in the message to the corresponding "value" field in the log output. The log output is as follows:

```
没有匹配到直播关键词  
[-emoji] { msgText: '我通过了你的朋友验证请求, 现在我们可以开始聊天了' }  
[-space] { msgText: '我通过了你的朋友验证请求, 现在我们可以开始聊天了' }  
[debug 号楼parser] ==> Time: 我通过了你的朋友验证请求, 现在我们可以开始聊天了  
[debug d-d点parser] ==> Time: 我通过了你的朋友验证请求, 现在我们可以开始聊天了  
[忽略 现在] ==> Time: 我通过了你的朋友验证请求, 我们可以开始聊天了  
maodou-nlp.js input before recognizeDateTime 我通过了你的朋友验证请求, 我们可以开始聊天了 +0ms  
maodou-nlp.js results after recognizeDateTime [ { time: null } ] +182ms  
[parser] ==> Time: { time: null }
```

对关键字的匹配与响应

- 会议模式
- 直播
- 课堂模式
- 提醒

```
const reg = /zoom|视频会议|音频会议|演讲|群学习/g

if(msgText.match(/直播/g)){
    let invite_url = '\n邀请连麦链接\nhttps://smh.maodou.io/invite/' + live_id
    let admin_url = '\n\n直播间后台链接\nhttps://smh.maodou.io/admin/content/course/live_id
    ....
}

else if(msgText.match(reg)){
    console.log(chalk.red('匹配到会议关键词'))
    let meeting_url
    if(msgText.match(/zoom|视频会议/g)){
        meeting_url = '\n视频会议链接\nhttps://kaihui.maodou.io/j/683175?mode=zoom'
    }
    else if(msgText.match(/音频会议/g)){
        meeting_url = '\n音频会议链接\nhttps://kaihui.maodou.io/j/683175?mode=audio'
    }
    else if(msgText.match(/演讲/g)){
        meeting_url = '\n演讲链接\nhttps://kaihui.maodou.io/j/683175?mode=lecture'
    }
    else if(msgText.match(/群学习/g)){
        meeting_url = '\n群学习链接\nhttps://kaihui.maodou.io/j/683175?mode=qunlearn'
    }
    ....
}
```



毛豆课堂小助手

明天下午6点半, bot5第五弹沙龙直播

[课程提醒创建成功通知]

标题: 弹沙龙
时间: 8/17/2019, 6:30:00 PM
地点: 半第五沙龙

消息原文:
明天下午6点半, bot5第五弹沙龙

毛豆课堂
弹沙龙

2019年07月
07/05 周五 钢琴课 18:00-19:00 (1/20)
07/06 周六 足球课 16:00-17:00 (1/10)
英语口语课 (1/10)

音频上传 视频上传 专栏 音频直播 视频直播

直播列表

序号	缩略图	名称	主讲人	状态	价格(元)	互动数	观看次数	分享链接	直播时间	排序	操作		
247		弹沙龙直播	毛豆小助手	即将直播	免费	0	0		2019/08/17 18:30				
246		拍照要打开	毛豆小助手	即将直播	免费	4	0		2019/08/09 10:42				
245		老师客户满意	毛豆小助手	即将直播	免费	0	0		2019/08/02 22:49				

毛豆课堂小助手

明天下午6点半, bot5第五弹沙龙演讲

[课程提醒创建成功通知]

标题: 弹沙龙演讲
时间: 8/17/2019, 6:30:00 PM
地点: 半第五沙龙

消息原文:
明天下午6点半, bot5第五弹沙龙演讲

演讲链接
<https://kaihui.maodou.io/j/683175?mode=lecture>

毛豆课堂小助手

明天下午6点半, bot5第五弹沙龙视频会议

[毛豆少儿课堂]
您已成功创建[
标题: 创建课程弹]
直播课

直播时间: 8/16/2019, 6:30:00 PM

直播地址: <https://maodouketang.com>
请您授权手机号并进入后台提前备课

温馨提示: 请您提前安排时间准时开课。

[你收到了一个小程序, 请在手机上查看]

凌晨

00:31

明天周五下午6点半, bot5第五弹沙龙群学习

[课程提醒创建成功通知]

标题: 弹沙龙群
时间: 8/16/2019, 6:30:00 PM
地点: 半第五沙龙群

消息原文:
明天周五下午6点半, bot5第五弹沙龙群学习

群学习链接
[https://kaihui.maodou.io/j/683175?
mode=qunlearn](https://kaihui.maodou.io/j/683175?mode=qunlearn)

周五明天下午6点半, bot5第五弹沙龙群学习

[课程提醒创建成功通知]

标题: 弹沙龙群
时间: 8/17/2019, 6:30:00 PM
地点: 半第五沙龙群

消息原文:
周五明天下午6点半, bot5第五弹沙龙群学习

群学习链接
[https://kaihui.maodou.io/j/683175?
mode=qunlearn](https://kaihui.maodou.io/j/683175?mode=qunlearn)

demo

- <https://maodouketang.com>

后续工作

- 小程序相关cdnthumb图片工作
- 聊天机器人训练引擎
- 讯飞语音stt
- 多提交PR

给wechaty 提交PR

- 本以为就我不会。。。

- 三个库有依赖关系， 提交顺序

- puppet-padpro

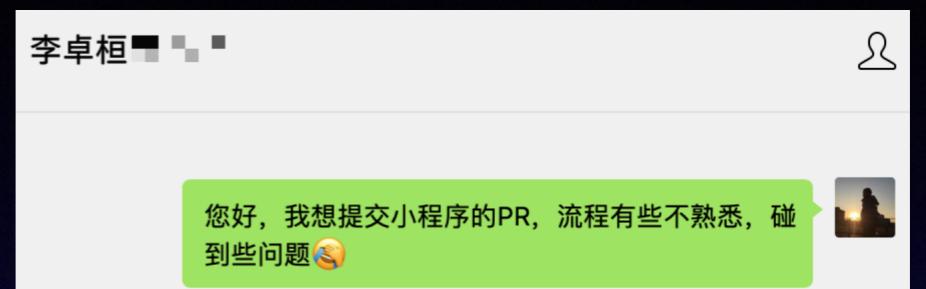
- puppet

- wechaty

- npm run lint 检查

- NO_HOOK=1 git push 避免版本号自动增加

- git 基本功



2019年7月19日 11:30



李卓桓

有个库，本地提交有版本号问题

```
chaty-puppet@0.15.24 lint:ts /Users/zhaoc/Github/wechaty-pr/wechaty-puppet
c --noEmit

chaty-puppet@0.15.24 lint:md /Users/zhaoc/Github/wechaty-pr/wechaty-puppet
mkdownlint README.md

5.25
ERR! code 128
ERR! Command failed: git tag v0.15.25 -am 0.15.25
ERR! fatal: tag 'v0.15.25' already exists
ERR!

ERR! A complete log of this run can be found in:
ERR!   /Users/zhaoc/.npm/_logs/2019-07-19T03_07_29_565Z-debug.log
ERR!   ed to exec pre-push hook script
ERR!   r: failed to push some refs to 'https://github.com/zhaoc/wechaty-puppet.git'
chaty-puppet git:(master) tig
<chaty-puppet git:(master) tig>
```

2019年7月19日 11:35

这样看来，你好需要先学习一下使用 git 的基本功能 tag

你可以设置一个环境变量 NO_HOOK=1 git push 试试。这样可以跳过一些检查。

谢谢大家